REPORT RESUMES

ED 003 619 56

ASSESSMENT OF A GRADUATE LEVEL SELF-TEACHING TECHNIQUE BASED ON A COMBINATION OF PROGRAMED INSTRUCTION AND THE CASE METHOD.

BY- WYNN, D. RICHARD MCKEEGAN, HUGH F.
PITTSBURGH UNIV., PA., SCHOOL OF EDUCATION

REPORT NUMBER NDEA-VIIA-939 PUB DATE SEP 63
GRANT OEG-7-48-1110-179
EDRS PRICE MF-$0.18 HC-$4.60 115P.

DESCRIPTORS- *CASE STUDIES (EDUCATION), *GRADUATE STUDY, *LAW INSTRUCTION, PROGRAMED INSTRUCTION, PROGRAMED MATERIALS, PROGRAMED UNITS, *TEACHING TECHNIQUES, ACHIEVEMENT TESTS, AUTOINSTRUCTIONAL METHODS, AUTOINSTRUCTIONAL PROGRAMS, *SCHOOL LAW, PITTSBURGH, PENNSYLVANIA

THIS STUDY ASSESSED THE VALUE OF A COMBINATION OF PROGRAMED INSTRUCTION AND THE CASE METHOD AS A GRADUATE-LEVEL TEACHING TECHNIQUE. TESTS, SYLLABI, AND PROGRAMED CASE STUDIES WERE DEVELOPED FOR CLASSES IN SCHOOL LAW AT FOUR UNIVERSITIES. THE STUDENTS WERE PRE- AND POST-TESTED AND WERE RATED ON BOTH ACHIEVEMENT AND AN OPINION SCALE. IN THE POST-ACHIEVEMENT AND RETENTION TESTS, THE EXPERIMENTAL GROUPS SCORED HIGHER THAN THOSE IN THE CONVENTIONAL CASE STUDY-DISCUSSION GROUPS. HOWEVER, NO DIFFERENCE COULD BE FOUND IN THE STUDENT ATTITUDES IN EITHER GROUP. THE CONCLUSIONS STATED THAT SUCH A PROGRAM APPEARED QUITE SUITABLE FOR LIMITED USE AS A GRADUATE-LEVEL TEACHING TECHNIQUE AND SHOULD PROVE SUITABLE IN THE AREAS OF MEDICINE, LAW, AND SOCIAL WORK. RECOMMENDATIONS FOR FURTHER STUDY SUGGESTED AN EXTENSION OF THE WORK FOR ADDITIONAL UNITS ON OTHER TOPICS AND A STUDY OF READING ABILITY AND ITS RELATIONSHIP TO ACHIEVEMENT AND STUDENT SATISFACTION WITH PROGRAMED MATERIALS. (PM)

ED003619

939 939

ASSESSMENT OF A GRADUATE LEVEL SELF-TEACHING TECHNIQUE BASED ON A COMBINATION OF PROGRAMMED INSTRUCTION AND THE CASE METHOD

D. RICHARD WYNN, ASSOCIATE DEAN, SCHOOL
OF EDUCATION, UNIVERSITY OF PITTSBURGH
HUGH F. MCKEEGAN, COORDINATOR OF EDUCATIONAL
MEDIA, DEPARTMENT OF PUBLIC INSTRUCTION,
COMMONWEALTH OF PENNSYLVANIA

The research reported herein was supported by a
grant from the U.S. Department of Health,
Education and Welfare--Office of Education
TITLE VII--PROJECT NO. 939
NATIONAL DEFENSE EDUCATION ACT OF 1958
GRANT NO. 7-48-1110-179

U.S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

SEPTEMBER 1963

Assessment of a Graduate Level Self-Teaching Technique Based on a Combination of Programmed Instruction and the Case Method

D. Richard Wynn
and
Hugh F. McKeegan

Department of Educational Administration
School of Education
University of Pittsburgh
September, 1963

Title VII Project Number 939
National Defense Education Act of 1958
Grant Number 7-48-1110-179

The Research Reported Herein was Supported
by a Grant from the U.S. DEPARTMENT OF HEALTH,
EDUCATION AND WELFARE -- Office of Education

TABLE OF CONTENTS

Page

Foreword

Acknowledgement

I. Introduction 1

A. Background of the Problem 1

B. Summary of Research 2

II. The Problem 5

A. Statement of the Problem 5

B. Elements of the Problem 5

C. Null Hypotheses 6

D. Source of the Data 6

E. Delimitations of the Problem 7

F. Limitations of the Research 7

III. Plan of the Study 8

A. Selecting Content Material for Programming 8

B. Constructing the Program 9

C. Constructing the Case Studies 12

D. Testing the Programmed Material and Case Studies . . 12

E. The Discussion Guide 20

F. Construction of the Evaluation Instruments 20

G. Preparing for the Experiment 22

H. Population for the Experiment 25

I. The Experiment 27

IV. The Data 32

A. Introduction 32

B. Primary Data 34

C. Secondary Data 54

V. Findings, Conclusions and Recommendations for Further Study 70

A. Findings 70

B. Conclusions 74

C. Recommendations for Further Study 78

Appendix A 80

Appendix B 87

Appendix C 90

Appendix D 101

Appendix E 104

Bibliography 108

Foreword

The educational programs in preparation for many of the professions have undergone a healthy re-examination and re-direction of goals, procedures and content in recent years. The preparation of future teachers and school administrators has not been overlooked in this process. So-called "five year" programs, cadet teacher programs, administrative interneships, campus residencies, the introduction of technological advances such as closed-circuit television and computer centered decision forcing devices are all examples of recent developments in the area of professional preparation in education.

This report summarizes an experiment in which one of the more successful traditional teaching methods, the case study approach, was combined with a newer innovation, programmed instruction, in the professional preparation of school administrators. The authors believe that the findings have implications not only for education but also for preparation in such fields as law, medicine, and accountancy. Further, the case study-programmed approach seems to be a likely method for continued in-service education in the professional fields. With the ever broadening scope of scientific and technological change, it is most essential that more adequate means of insuring continued competency in every field be developed. It is hoped that this report will contribute in some small way to this end.

This research had its inception as a proposal for a doctoral study at the University of Pittsburgh. Preliminary guidelines for the research were developed under the guidance of the late Robert W. Brittell, Professor of Education. Subsequently, the programmed material was developed, the experiment conducted, and this report prepared by the second author with the assistance and supervision of the first. The many other professional colleagues who cooperated in this work are recognized in the Acknowledgement which follows.

Acknowledgment

The following members of the staff of the University of Pittsburgh assisted with this project. Dr. C. Mauritz Lindvall provided help in the development of the programmed material and the evaluative instruments. Dr. J. William Asher provided expert guidance regarding the overall research design and statistical procedures. The manuscript from which this report was prepared was critically reviewed by Dr. Donald L. Cleland and Dr. Maurice J. Thomas.

Dr. J. William Moore of Bucknell University served as programming consultant and gave valuable suggestions as to research procedures.

Mr. Jack Quackenbush of the Pennsylvania State University arranged for the testing necessary in the development of the attitude scale.

Dr. Lee O. Garber of the University of Pennsylvania, Dr. James J. Jones of Temple University, Dr. Franklin Miller of the Pennsylvania State University, and Dr. Neal V. Musmanno of the University of Pittsburgh not only permitted the experimental material to be used with their classes, but gave of their own time and advice to insure the success of the project.

Mr. Emmanuel Berger of the Research Bureau, Pennsylvania Department of Public Instruction, assisted with the statistical analysis of the data.

Dr. Paul Masoner, Dean of the School of Education of the University of Pittsburgh, and Dr. Charles H. Boehm, Superintendent of Public Instruction, Commonwealth of Pennsylvania, facilitated the work of this study in every way possible. Dr. Thomas Clemens and Dr. Kalmar Stordahl of the U.S. Office of Education gave freely of their time and advice whenever called upon.

Finally, the foresight of those members of the Congress and Senate of the United States who, through the National Defense Education Act, provided funds for research projects such as this must not go unrecognized.

The authors wish to acknowledge the assistance of all the people mentioned above and to extend to them a most sincere "Thank you" for their efforts.

I. Introduction

A. Background of the problem

The technique of breaking down material to be learned into small, incremental steps, arranging these steps so that student response is required and providing the student with immediate knowledge of results; i.e. the technique called programming, seems to offer tremendous possibilities for the improvement of the instructional process on all levels.

This technique, which can in part be traced back to the pioneer work done by Sidney Pressey (1926) and his associates at Ohio State University in the 1920's, received new support and a new theoretical framework from Dr. B. F. Skinner's now famous paper, "The Science of Learning and the Art of Teaching" (1954.) In this article, Dr. Skinner outlined how operant conditioning principles developed by the experimental psychologist in laboratory studies might well be applied to verbal material taught in the school situation. Dr. Skinner's emphasis on the "program"; on the careful arranging and sequencing of items or frames so as to shape the desired behavior, and on the necessity for constructed responses distinguish his work from that of Pressey.

Since 1954, rather extensive programs have been developed to teach a variety of subject matter on all levels from the primary grades through college subjects. Intensive work is being done, particularly in the areas of elementary and secondary mathematics and sciences. On the other hand, very little programming has been completed in the social science area, although there are indications of extensive developmental work in this field.

With the exception of certain experimental work in medical education, little has been done in programming graduate level work in the professional schools.

Beck (1963) comments as follows "...we have overlooked what may be the most potent use of programmed instruction: the training of teachers." He then goes on to suggest that the "staples" in teacher education might well be prepared in programmed form and that students be required to complete these individually so that professional courses could be limited to higher-level seminar type offerings.

Schramm (1962) also suggests that programmed instruction should be "a particularly powerful tool in teacher education..."

However, much more is involved than investigations of whether programmed instruction can be used in teacher and graduate education. Industry, the military, and business are already making extensive use of programs for training purposes. While the purposes and objectives differ, certainly some of this material might be used in appropriate graduate courses. Many questions remain, however, concerning the efficiency of programming on this level, both as an isolated technique, and in combination with other techniques in a "systems" approach. The parameters of programmed instruction also require further exploration for this population. Manner of presentation, mode of response, size of step, type of reinforcement, transfer values are but a few of the variables worthy of further study.

B. Summary of Research

A fairly large number of evaluative type studies conducted over a period of three decades indicate that programmed materials in a variety of forms are quite useful in instruction. The fact that programmed materials can "teach" has been well established. That they also reduce variability in student performance has been a finding of several researchers. Whether this phenomenon is due to characteristics of programming itself or to limitations of the instruments used for measurement is an unsettled point.

The contribution to efficient learning of such programming principles as providing immediate knowledge of results, proper sequencing of material to be learned, and arranging for response to stimuli has been rather thoroughly investigated. How best to apply and combine these principles in program design is a largely unexplored area.

Practically all of the research in programming to date, whether evaluative or parametric in nature, has been based on the use of programmed material alone rather than in combination with other instructional approaches. An obvious next step is to assess how programmed material can best be combined with other instructional methods and what effect those combinations would have on such factors as efficiency, achievement and retention.

The whole area of the relationship of student attitude to acceptance of and performance with programmed material requires further research. Practically all of the studies touching on this point indicate that attitude doesn't matter as far as performance is concerned. These findings are usually based on very brief exposures to programmed materials. In most cases the validity and reliability of the attitude measuring instruments used leave much to be desired.

Among the parametric studies such factors as size of step, constructed versus selected response, pacing, branching, and mode of response have been studied. In comparisons of overt versus covert response, a "no difference" finding in achievement with greater efficiency time wise for covert response is frequently reported. (Evans, 1960; Silberman and Alter, 1961; Walker and Stolurow, 1962). Goldbeck (1960) found response mode related to both completion time and item difficulty.

In regard to response made studies, the question is raised whether significant differences would be found with the time factor held constant through requiring non-overt groups to go through the program again, for example. Since in this situation non-overt response would be more efficient time wise, the additional time secured might produce significant differences in learning.

The values of combining programmed materials with other types of instructional procedures also merits investigation.

II. The Problem

A. Statement of the problem

The purpose of this study was to assess the value of a combination of programmed instruction and the case method as a graduate level teaching technique and further to investigate the values of certain variations in mode of response within this setting.

B. Elements of the problem

The elements of this problem included the following:

1. To assess in terms of student achievement the comparative effectiveness of the case study-programmed overt response approach, the case study-programmed covert response approach, and the case study-discussion approach.

2. To assess in terms of student retention the comparative effectiveness of the case study-programmed overt response approach, the case study-programmed covert response approach, and the case study-discussion approach.

3. To determine the relation between measures of student opinion of a particular approach and achievement in that approach.

4. To determine the relation between measures of student opinion of a particular approach and retention of material taught via that approach.

5. To attempt to determine, via personal interview techniques, additional factors that may be concomitants of:

   a. high and low achievement.

   b. a high degree or a low degree of retention.

   c. a strong preference or lack of a preference for a particular approach.

C. Null Hypotheses

The following four null hypotheses were established as a basis for statistical analysis of the data. Findings regarding the first four elements of the problem were based on acceptance or rejection of these hypotheses:

1. No difference in achievement will be found among the case study-programmed overt response approach, the case study-programmed covert response approach and the case study-discussion approach.

2. No difference in retention will be found among the three experimental groups.

3. No relationship will be found between student opinion of any teaching approach under study and achievement in that approach.

4. No relationship will be found between student opinions of any of the three approaches and retention.

D. Source of the data

The fall 1962 enrollees in the graduate courses in school law at the University of Pittsburgh, The Pennsylvania State University, The University of Pennsylvania, and Temple University were the subjects who provided the primary source of data for this experiment. The data collected included the following:

1. Subjects' scores on the achievement, retention, and opinion measuring instruments.

2. Personal data regarding all subjects gathered as part of the written instruments.

3. Written responses to the program by the case study-programmed overt response group.

4. Tape recorded interviews with selected subjects.

5. Tape recordings of the discussion sessions.

Secondary sources of data included the following:

1. Texts, references, and researches in School Law.

2. Texts, references, and researches in programmed instruction and programming techniques.

3. Interviews and correspondence with the various instructors in school law and those skilled in programming theory and techniques.

E. Delimitations of the problem

The case study programmed material developed for this research was based on a selected area of content in the graduate course in school law. This content area is generally titled "Federal Relations to Education."

The experimental material was equivalent to approximately 185 minutes of instruction time. It contained 209 frames requiring a total of 443 responses. Nine case studies of major importance to the topic were integrated with the program.

The data collected in this research were based on the experimental use of this material with school law classes in the four large Pennsylvania universities.

F. Limitations of the research

This research was limited by the rather short exposure of the various groups to the experimental treatments. Findings should be interpreted with this condition in mind.

Further, the findings of this research were obtained using programmed material in conjunction with case study material. Caution should be exercised in generalizing from these findings to findings based on the use of programmed material alone.

III. Plan of the Study

A. Selecting content material for programming

The professors of school law at the four universities involved were interviewed in December 1961. In these interviews their consent to participate in the study was secured and their opinions were sought as to the areas of the course suitable to the case study-programming approach.

Two of the 4 instructors preferred a topic or topics ordinarily presented in the early part of the course. Two instructors also stated that they could grant no more than 200 minutes of class time for experimental purposes. Time was a particularly critical factor at one school, operating on the quarter system, where the law classes met for just fourteen 160 minute sessions. At this school the instructor felt that he could give up no more than one class session for the study. However, the students were asked to come early so that the required 200 minutes instruction time might be reached.

Exploration of possible subject matter areas for the study led to the conclusion that "Federal Relations to Education" and "State Relations to Education" seemed most amenable to the method of the study and the criteria of time and placement in the course content established by the instructors.

Texts, syllabi, and other references were then reviewed and a tentative outline of the content covered in the topic "Federal Relations to Education" was developed. This outline was designed so that the various instructors might indicate points emphasized, points given minor emphasis, and points not included in their coverage of this topic.

The professors at the 4 schools were interviewed again in April and May 1962 and their reactions to the outline were secured. All the instructors approved the outline with minor revisions and additions. As the programmed material and case studies were developed, it became evident that it would be impossible to cover the entire outline in the time allotted for the experiment. Consequently, certain topics included in the original outline had to be eliminated from the program finally used in the experiment.

B. Constructing the program

The first step in constructing the program involved describing the specific verbal skills the student should have mastered at the conclusion of each section of the program. This followed the standard practice of programmers in first identifying the terminal behaviors being sought. Some difficulties were encountered at first in stating these objectives in student centered rather than content centered terms.

Following programming practices of J. William Moore at Bucknell University and others, these objectives were then stated as "criterion" items or "test" items for inclusion in the program. These criterion items contained no cues, hints, or other instructional aids. They were designed as test items and identified as "C" items (criterion items) in the program. A series of instructional items was developed for each criterion item.

These instructional items were designed to "teach," with the first few items in each series containing liberal hints and cues to insure a high degree of correct response. The hints and cues were diminished as the instructional items lead up to each criterion item.

Since good programming is good teaching, and since the skilled teacher reviews material as the needs of students may require, review items were developed and inserted in the program. These review items included restatements of criterion items, major facts regarding case studies and summarization items at the conclusion of each unit. A typical series of items is shown in the following table.

TABLE 1

TYPICAL SEQUENCE OF PROGRAM ITEMS

1. Instructional Item
2. " "
3. " "
4. " "
5. " "
C- 6. Criterion Item
7. Instructional Item
8. " "
9. " "
10. " "
11. " "
R-12. Review Item
C-13. Criterion Item
14. Instructional Item
15. " "
16. " "
17. " "
R-18. Review Item
C-19. Criterion Item
20. Instructional Item
21. " "
22. " "
C-23. Criterion Item
R-24. Review Item
R-25. " "

Using sequences such as this permitted the researcher to identify rather specifically those areas where the program "broke down" or "failed" to teach. If, for example, a subject responds correctly to criterion item C-6 but misses its corresponding review item, R-12, then the programmer knows he must insert additional review items somewhere between C-6 and R-12. If the subject misses C-13 then the programmer must revise instructional items 7 through 11 or add additional items to bring the subject to criterion level performance. Should the subject respond correctly to items C-6, R-12, C-13, and R-18 but then miss the same items re-stated in R-24 and 25, the insertion of additional review items is again indicated.

Use of such a programming system has several advantages; it enables the programmer to pin-point weaknesses in a program as it is developed; it discourages over-teaching and the use of an excessive number of frames to shape-up a behavior; it emphasizes to the student what he must know, and finally it places error rate in proper perspective. A low error rate in the instructional items may be desirable but it is not absolutely essential. Error rate in the criterion and review items is of the utmost consequence, however, since it is an indication of the real instructional efficiency of the program.

The program used in this experiment incorporated legal case studies in its format. Frames 1 through 25 in the illustration above, for example, might have taught those legal concepts necessary to an understanding of and interpretation of the Jones v. Smith case (hypothetical.) The student was then referred to this case study itself. The case study would be followed by a series of frames designed to insure understanding of the case and its implications. All of the frames in the program followed the constructed response format. However, many of the post-case study frames resembled Pressey type items in that they were not self contained but asked for responses based on the case readings.

The final program used in the experiment contained 5 topics and a general review. A total of 9 case studies were incorporated in the topics. Table 2 is a list of the topics and case studies covered by the program.

A further analysis of the final program by type and number of responses is presented in Table 3.

C. Constructing the case studies

Summaries of the case studies to be used with the program were prepared from standard law library references, mainly the National Reporter System and from related legal references. Clark Spurlock's work (1955) proved very valuable in this regard.

Each case study included background material, the facts of the case at issue before the Court, and important excerpts from the Court opinion. The 9 case studies combined totalled 32 pages of reading material. In final form they ranged in length from 543 words to 1,712 words with a median length of 940 words. A sample case study is included in Appendix B.

D. Testing the programmed material and case studies

1. Individual Trials - As each section of the programmed material was prepared it was tested on individual subjects. Items were prepared on 5 x 8 cards with the item on one side of the card and the correct response on the other. Subjects were asked to read each item, record their answer on an answer sheet, check the correctness of the response and proceed to the next item. At the conclusion of each trial the experimenter analyzed the answer sheets and revised the program by inserting, deleting or revising frames.

2. Pilot Trial - The first 3 units were given a group trial in the 1962 summer session Law class at Temple University. A total of 9 graduate students participated. So that each student might have his own set of materials, the program was typed 5 frames on a sheet on ditto sheets and duplicated. Each sheet was then cut into 5 parts so that each frame was on a separate sheet of paper approximately $2\frac{1}{4}$" x $8\frac{1}{2}$" in size. The frames were then arranged in order and stapled by units. The responses had been

TABLE 2

SUMMARY OF CASE STUDY PROGRAMMED MATERIAL BY TOPICS AND COURT CASES INCLUDED

Topic	Title	Case Studies
A	Education and the Constitution	None
B	Education and the General Welfare	1. U.S. v. Butler, 297 U.S. 1 (1936)
		2. Helvering v. Davis, 301 U.S. 619
C	Contracts, Education and the Constitution	3. Trustees of Dartmouth College v. Woodward, 4 Wheaton (U.S.) 518 (1819)
D	The Fourteenth Amendment and the Police Power	4. Meyer v. Nebraska, 262 U.S. 390 (1923)
		5. Pierce v. Society of the Sisters of the Holy Name of Jesus and Mary, 268 U.S. 510 (1925)
		6. Cochran v. Louisiana, 281 U.S. 370 (1930)
E	Education, Religion, and the Constitution	7. Everson v. Board of Education, 330 U.S. 1 (1947)
		8. Illinois ex. rel. v. McCollum, 333 U.S. 203 (1948)
		9. Zorach v. Clauson, 343 U.S. 306 (1952)
	General Review	None

TABLE 3

ANALYSIS OF PROGRAM BY TYPE OF ITEM AND NUMBER OF RESPONSES REQUIRED

	Type of Item						Total	
	Instructional		Criterion		Review			
	No. of Items	No. of Responses	No. of Items	No. of Responses	No. of Items	No. of Responses	No. of Items	No. of Responses
Topic A	55	96	8	25	5	19	68	140
" B	43	74	8	21	3	13	54	108
" D	17	31	5	12	6	18	28	61
" E	27	45	11	30	5	18	43	93
" C	8	15	2	5	--	--	10	20
Review	--	--	--	--	6	21	6	21
TOTAL ENTIRE PROGRAM	150	261	34	93	25	89	209	443

typed so that the responses to frame 1 appeared at the top of frame 2, the responses to frame 2 at the top of frame 3, etc.

The case studies required for the first three units were also prepared in ditto form.

Although the final evaluation and opinion instruments had not yet been developed, since the programming was not yet complete, a test and several opinion statements were developed for use with this group.

Students took the pre-test and began working with the programmed material and case studies on July 2. The material was completed and the post-test and opinion statements administered on July 3. Several students were interviewed to obtain additional information regarding their reactions to the material.

Individual error rates for the programmed material used in the pilot study ranged from .053 to .322 with a mean error rate of .156. The results of this pilot trial indicated the need for further revision of the program. In this regard, it should be noted that two subjects accounted for 44% of the total errors made by the group.

An attempt was made to keep a record of the time taken by each subject to complete each unit. Two subjects failed to follow instructions completely and their time for certain sections had to be estimated.

The following table summarizes the time data:

TABLE 4

RANGE AND MEAN TIME IN MINUTES

PILOT TRIAL - TEMPLE UNIVERSITY

7-2-62 7-3-62

	Topic A	Topic B	Topic C	Total Program
Range (in minutes)	18-38	25-45	45-70	95-180
Mean time (in minutes)	23.42	34.85	58.88	129.55

The need for further refinement of the program and case study material to remain within the time allotments at the various schools was clearly indicated.

The survey of student opinions indicated generally favorable attitudes toward the programming-case study approach. The opinion statements used in the pilot trial are listed in Table 5 together with the number of responses in each category. One student of the nine in the group failed to return these statements.

Four of the eight subjects volunteered additional comments in writing and follow-up interviews were conducted with seven of the nine participants. The oral and written comments indicated dissatisfaction with the format of the material and the extended use in class. Reactions to the programming itself were generally satisfactory.

TABLE 5

SURVEY OF OPINIONS REGARDING CASE STUDY-PROGRAMMED MATERIAL IN SCHOOL LAW

PILOT GROUP - TEMPLE UNIVERSITY

	1 Strongly Disagree	2 Disagree	3 Undecided	4 Agree	5 Strongly Agree
1. There is no place for material of this type in graduate level work.	2	4	2		
2. I would prefer this material to a conventional text.		1	1	5	1
3. This material is better suited to individual use outside of class time.			1	5	2
4. Class time should never be spent studying material such as this.	1	4	1	1	1
5. I missed the "give and take" of class discussion while using these materials.		1	2	3	2
6. Requiring overt response to each item seems unnecessary at this level.	1	5	1	1	
7. I learned more from the case reading than from the programmed material.	1	3	1	2	1
8. The case readings didn't add much to the learning situation.	3	5			
9. Most of the frames or steps in the program were large enough to be challenging.	1	1	2	4	
10. The programmed material contained too many one-word responses.	1	5		1	1

A 28 item achievement instrument was prepared for use in this pilot trial. This test was used both as a prior test of student knowledge of the topic and as an achievement test at the conclusion of the program. The mean score for the pre-test, based on an N of 8, was 8.12 and for the post-test 17.87.

Having computed the standard error of the difference between the means for paired observations, the value of t was determined as 5.90. With 7 degrees of freedom, the obtained value exceeded the score of 5.405 required at the .001 level (two-tailed test) and well exceeded the probability for a one-tailed test of significance (2.702 at .001 level.) The null hypothesis was therefore rejected and the conclusion drawn that the introduction of the programmed-case study materials in the interval between pre and post tests did produce a significant positive change in achievement.

Because of the preliminary nature of the instruments used in the pilot trial and small number of subjects involved, no further statistical analysis of the data was attempted. However, the following conclusions were drawn regarding further development of the project.

a. Additional revisions of the programmed material were necessary in order to reduce the erratic error rate.

b. The programmed material and case studies particularly required editing to permit the principal experiment to be carried out within the allotted time limits.

c. Observation of the students during the experiment indicated that the format of the material caused much frustration and loss of time. The survey of student opinions and further personal interviews emphasized student dissatisfaction in this regard.

d. Despite the deficiencies in the material, there was strong evidence based on preliminary statistical analysis of achievement and on student opinions that the program did in fact "teach" and that these graduate students were favorably impressed by its content.

3. Further revision and development - An item analysis of the programs used in the pilot study was conducted; every incorrect response for each item was listed and further revisions were made. Units D, E, F, were developed and revised using individual subjects from summer classes at the University of Pittsburgh and professional staff members of the Department of Public Instruction. Following are the error rates of the subjects used in the last 5 revisions of the program.

TABLE 6

INDIVIDUAL ERROR RATES FOR FINAL PROGRAM REVISIONS

Subject	Unit A	Unit B	Unit C	Unit D	Unit E	Unit F
S	.050	.040	.057	.072	.051	.070
T	.033	.070	.034	.217	.025	.011
U	.190	.140	.125	.114	.135	.148
V	.080	.117	.159	.057	.135	.044
W	.059	.057	.092	.063	.060	.070

A study of the time consumed by various subjects indicated that the program and case studies would require 225 to 275 minutes total time. After finding it impossible to secure additional time at 2 of the 4 schools, it was decided to drop the original Unit A dealing with the topic of "sovereignty" entirely from the program and further edit the case studies. The final content of the experimental material has already been summarized in Table 2.

The final version of the program was duplicated in vertical format on 8½ x 11 paper, a plain sheet of blue paper was inserted between each program sheet and each program was punched and bound, using continuous ring plastic binders. Each blue sheet was then cut to a width of approximately 2¼". Each of these strips was again divided so that each response was hidden by an individual hinged cover.

The case studies were duplicated in order on 8½ x 11 paper and stapled. A set of these case studies was provided with each program and for each participant in the discussion group.

E. The discussion guide

An outline of the major topics covered in the programmed material was prepared for use with the discussion group. This outline included the same chart used in the program to illustrate those sections of the Federal Constitution pertaining to education. The discussion guide is included in Appendix E.

F. Construction of the evaluation instruments

1. The opinion instrument - The opinion instrument consisted of two parts: a 16 item Likert type scale designed to measure student attitude towards the teaching approach to which they had been exposed, and 19 multiple choice items arranged to secure opinions regarding specific aspects of the experiment.

The Likert Scale was constructed in the following manner. Over 40 items all centered on the dimension of teaching method were drawn up. These were then reviewed by two people competent in this field, revised, and used with several individual subjects. As a result of these reviews and trials the original list of statements was revised and reduced in number from 44 to 36. Twenty-three of these statements were positive in nature and 13 negative in nature.

Arrangements were made to further refine these items using groups of students who were receiving instruction via programmed material. Two such groups were located at The Pennsylvania State University. They were college freshmen participating in an experiment designed to assess the value of television and teaching machines as presentation modes. Three sections (N = 40) were receiving programmed instruction in English grammar over a closed circuit television system. The fourth section (N = 20) was studying the same programmed material presented by a simple teaching machine. The attitude scale was administered to all 4 groups as part of their regular class work.

The scales were scored using standard procedures outlined by Goode and Hatt (1952, pp. 275-276). The upper and lower quartiles were identified in each group and the Discriminating Power of each statement determined by finding the difference between the mean scores for each statement within these quartiles.

Both negative and positive statements were also ranked as to discriminating power within each group. On the basis of this ranking and discriminating power, 8 positive and 8 negative statements were selected for inclusion in the final scale. These 16 statements were of course re-arranged in order and re-numbered in the final scale.

The remainder of the opinion instrument consisted of 19 multiple choice items designed to assess student opinions regarding the case studies, the programmed material, potential uses of the programmed material and the discussion approach. A copy of the instrument may be found in Appendix C.

2. The achievement instrument - The final achievement instrument contained 52 items in its first draft. However, when one unit of the

program had to be eliminated due to restrictions of time, the instrument was shortened correspondingly. In final form it contained 40 items, all of the multiple choice type. Multiple choice items were selected because they are generally conceded to be superior objective type items and because they represented a minimum carry-over from the completion items used in the program.

Face validity was assured by constructing the multiple choice items on the basis of what had been taught in the criterion items of the program.

Estimates of reliability were difficult to obtain in advance of the experiment since no group of sufficient size, having the time to study the program and take the achievement test, could be found.

The 52 item version of the test was administered to 8 graduate students enrolled in a summer course in school law at the University of Pittsburgh. Their comments both written and oral helped to eliminate poorly worded or tricky items.

Sample items from the achievement instrument can be found in Appendix D. The same instrument was used with all groups for pre-testing and for tests of immediate achievement and retention.

G. Preparing for the experiment

1. The following materials were also prepared or secured for the experiment.

a. A personal data card - This card was designed to secure pertinent data regarding such factors as number of graduate credits completed, extent of teaching and/or administrative experience, and undergraduate majors and minors.

b. Written instructions - These were an exact duplicate of the instructions included in the program. They were duplicated separately to permit distribution of these instructions in advance of the use of the program.

c. Experimenter's remarks - To assure uniformity at each school, the experimenter's introductory remarks and oral directions were prepared in written forms. With minor editing to suit each situation the same remarks and instructions were given at each school.

d. Tape recorder and tapes - To further assure uniformity the instructions given at each school together with any questions raised by students were recorded for possible future analysis. During the experiment itself the proceedings of the case study-discussion group were recorded for further study.

e. Proctor's instructions - Written instructions were prepared to guide the proctors who would supervise the programmed overt and programmed covert groups at each school.

2. The professors of school law at the four participating universities were contacted to establish the exact schedule for the experiment. This Schedule is summarized in Table 7.

TABLE 7

SCHEDULE ESTABLISHED FOR THE EXPERIMENT AT THE VARIOUS UNIVERSITIES

	University			
	A	B	C	D
30 Min. Introduction Oral Directions Pre-test	Thurs., Sept. 13 (9:30-10:00 P.M.) (2nd class)	Sat., Sept. 15 (9:30-10:00 A.M.) (1st class)	Tues., Sept. 18 (8:40-9:10 P.M.) (1st class)	Fri., Sept. 28 (7:15-7:45 P.M.) (1st class)
180-195 Min. Experiment & Post-test	Thurs., Sept. 20 (6:30-9:45 P.M.)	Sat., Sept. 22 (9:15-10:50 A.M.) (11:20-1:00 P.M.)	Tues., Sept. 25 (7:55-9:20 P.M.) Tues., Oct. 2 (7:50-9:30 P.M.)	Fri., Oct. 5 (6:55-10:05 P.M.)
Distribution of Opinion Instrument	September 20	September 22	October 2	October 5
Retention Test	October 18 (6:45-7:00 P.M.)	October 27 (9:15-9:30 A.M.)	October 30 (7:55-8:10 P.M.)	November 9 (7:15-7:30 P.M.)
Elapsed Time Post-test - Retention Test	4 weeks	5 weeks	4-5 weeks	5 weeks

NOTE: The times indicated in parentheses under each heading are the actual times spent in the various phases of the experiment at each school.

H. Population for the experiment

In three of the four universities the experimental group consisted of the enrollees in the graduate course in "School Law." The title of this course varied slightly from university to university as did the credit hour requirements. In the fourth university the enrollees in a graduate course in educational administration, "Introduction to Educational Administration," made up the experimental group. Four meetings of this class were devoted to a survey of school law taught by the school law instructor. Other major topics in the administration area were surveyed in a similar manner.

In all the schools but the latter the enrollees received instruction in school law from their regular instructors during the interim period between the immediate achievement test and the retention test. The "Introduction to Educational Administration" group received no instruction in school law during the interval between tests.

The following tables summarize the distribution of subjects within the various schools on those variables for which data was available.

TABLE 8

DISTRIBUTION OF SUBJECTS WITHIN UNIVERSITIES BY YEARS OF TEACHING EXPERIENCE

Years Experience	University A	B	C	D	Total
0	1			1	2
1- 3	3	11	5	4	23
4- 6	1	5	13	7	26
7- 9	5	2	13	3	23
10-12	2	3	5	1	11
13-15	4	2	2	2	10
16-18		1	1	1	3
19-21			1	1	2
22-24			1		1
25 +		1			1
Total	16	25	41	20	102

TABLE 9

DISTRIBUTION OF SUBJECTS WITHIN UNIVERSITIES BY NUMBER OF EARNED GRADUATE CREDITS

Graduate Credits Earned	University A	B	C	D	Total
0- 9	2	5	2		9
10-19		8	4		12
20-29		3	5	4	12
30-39	1	3	13	3	20
40-49	3	1	6	6	16
50-59	2	3	5	3	13
60-69	3	1	2	1	7
70-79	3	1	1	2	7
80-89	1		2	1	4
90-99			1		1
100 or more	1				1
Total	16	25	41	20	102

TABLE 10

DISTRIBUTION OF SUBJECTS WITHIN UNIVERSITIES BY YEARS OF ADMINISTRATIVE EXPERIENCE

Years Experience	University A	B	C	D	Total
0	10	19	26	16	71
1- 3	1	5	5	1	12
4- 6	4	1	8	1	14
7- 9	1		1		2
10-12			1	2	3
Total	16	25	41	20	102

TABLE 11

DISTRIBUTION OF SUBJECTS WITHIN UNIVERSITIES BY SEX

	University A	B	C	D	Total
Male	15	18	36	18	87
Female	1	7	5	2	15
Total	16	25	41	20	102

I. The experiment

1. First session - The first session at each school was spent in giving the necessary instructions to the class, filling out the personal data cards, administering the pre-test, and randomly assigning each student to one of the three treatment groups.

a. Instructions - At each school the experimenter used basically the same set of oral instructions. These were prepared in written form in advance. A duplicate set of the written instructions included in the program was distributed to all students to orient them further to the project.

b. Data cards and pre-test - After the instructions were completed each student filled out the data card and responded to the pre-test under the experimenter's supervision. The regular instructor at each school also assisted in the administration of these instruments. Administration of the testing instrument took no more than 15 minutes. A summary of the results of the administration of this test is presented in Table 12.

TABLE 12

RANGE, MEAN, AND MEDIAN OF PRE-TEST SCORES BY TREATMENT GROUPS

	Group		
	Overt (N=35)	Covert (N=35)	Discussion (N=32)
Range	2-23	4-19	2-26
Mean	12.51	11.60	12.78
Median	13.33	11.75	12.75

After the pre-tests had been collected the members of the discussion group were notified to report to a separate designated room for the next session. Remaining members of the class (the overt response and covert response groups) were asked to report to their regular classroom at the next session. No indication of assignment to overt or covert groups was given in advance of the second session.

2. Second session

a. All groups - In preparation for the second session a kit of materials was prepared for each participant. These materials were enclosed in a large manilla envelope marked with the individual's name. Members of the overt group at each school received a "blue" program and a set of case studies, members of the covert group received a "green" program and a set of case studies, and members of the discussion group received a discussion guide and a set of case studies.

b. Overt and Covert groups - Arrangements were made to have a proctor supervise the overt and covert groups at each school. These proctors were given written and oral instructions in advance regarding their role in the experiment. In the oral instructions it was emphasized that they were to give no help of an instructional nature to the subjects. Their duties included distribution of materials, a review of the instructions given the previous week, and general supervision of the groups as they worked with the materials. At 50 minute intervals proctors were to announce the time and ask subjects to record this time at whatever point they were working in the program or case studies.

To help assure most efficient use of the time allotment at each school, Subjects were asked to study the topics in the following order: "A," "B," "D," "E," "C," "General Review." "C" is a very short unit containing only 1 case study. As far as can be determined from the time checks all members of the covert group (N=33) followed these instructions. However, 3 members of the overt group (N=28) either failed to follow these instructions or recorded their responses improperly. In addition the size of the groups was reduced by absences as follows: overt group -- 3 absent, covert group -- 2 absent, discussion group -- 1 absent. Four of the 6 absences occured at the school where this second session had to be scheduled on a split period basis. Due to these absences and failure to follow instructions the N's for the overt, covert and discussion groups were reduced to 28, 33, and 31 respectively.

Due to variations in time patterns and the efficiency of proctors in expediting the work, it was impossible to completely eliminate any differences in time spent by the groups among the various schools. However time was equated for all subjects at each school, regardless of the treatment group to which they were assigned. Among the various schools the actual time spent in this second session, including the immediate achievement test, varied from 185 to 195 minutes. No more than 15 minutes was spent in testing. Proctors distributed the opinion instruments, enclosed in return-addressed, stamped envelopes, as the test was completed.

c. Discussion group - The experimenter lead the discussion groups at each school. The discussions were tape recorded for further analysis. In using the discussion technique, an attempt was made to parallel the activities of the programming groups as to content and time expended on each topic. The post-test was administered under the supervision of the experimenter and the opinion instrument distributed as the test was completed. All experimental materials used by the overt, covert, and discussion groups were collected at the end of this session.

3. Third session - The experimenter re-visited all four schools after an interval of 4 to 5 weeks and administered the testing instrument again as a measure of retention. A total of 5 students who had participated in the first and second sessions were absent at the time the retention test was given. One of these students was from the overt group, and 2 each from the covert and discussion groups. Thus the size of the groups for retention measurement purposes was reduced to 27, 31, and 29 respectively.

4. Individual Interviews

a. Selecting subjects for interviews - All subjects regardless of treatment group were ranked from high to low on both the achievement instrument and the opinion scale. Individuals having extremely high or low rankings on either instrument were selected for further interviews. A total of 22 participants were interviewed; 10 personally and the remainder by telephone.

IV. The Data

A. Introduction

1. Reliability of the testing instrument - Reliability of the test was determined by application of the split-half reliability technique to the immediate achievement scores of the combined groups. Each subject's test results were divided on an odd item - even item basis and the Pearson product-moment correlation coefficient calculated. This coefficient indicated a correlation of.77 between the two halves of the test.

To determine the estimated reliability coefficient of the entire test the Spearman-Brown prophecy formula was applied, yielding a reliability coefficient of .87 for the entire test.

2. Reliability of the opinion scale - Procedures similar to those outlined for the testing instrument were followed with the opinion scale. Each subject's total score on the 16 item scale was divided on an odd - even basis, the correlation coefficient calculated and the reliability determined according to the Spearman-Brown prophecy formula. This procedure yielded a reliability of .94 for this instrument.

Both total program error rates and criterion item error rates are reported in the tables which follow. In addition, error rates for those items classified as review items and for review and criterion items in combination are reported.

In this regard criterion item error rate is probably the better index of program quality. A high error rate on criterion items is a definite indication that the program is not achieving its objectives. Failures in this regard are sometimes masked when only total error rates are reported.

All of the error rates reported are, of course, based on the answers recorded by the overt response group. In programs of this type it is impossible to list all possible synonyms for the correct response. In determining the error rates, each program was checked individually and equivalent or similar responses accepted as correct.

TABLE 13

TOTAL PROGRAM ERROR RATE

(OVERT GROUP ONLY)

Topic	Total No. Responses	Incorrect Responses	Error Rate
A	4,620	283	.061
B	4,019	183	.045
D	1,952	123	.063
E	2,791	156	.055
C	480	28	.058
Gen. Rev.	546	75	.137
GRAND TOTAL	14,408	848	.058

TABLE 14

ERROR RATES FOR CRITERION ITEMS, REVIEW ITEMS AND CRITERION AND REVIEW ITEMS COMBINED

Topic	Criterion Item Error Rates	Review Item Error Rates	Combined Criterion and Review Item Error Rates
A	.056	.104	.079
B	.045	.080	.058
D	.044	.099	.077
E	.057	.071	.062
C	.002	----	.002
Gen. Rev.	----	.137	.137

The review item error rate is to a large extent a function of the number of review items inserted in the program. In both the criterion item error rate and the criterion-review item error rate the program used in this experiment reached generally acceptable levels.

B. Primary data

In this section the data and statistical procedures con~~dis~~sidered essential to the elements of the problem are reported. The .05 level of confidence is the basis for rejection of the null hypotheses. Differences at the .10 level will be called to the reader's attention. Differences not reaching the .10 level are indicated by the symbol (N.S.) in the appropriate tables.

1. Instructional approach and achievement - Data secured from the pre-test and achievement test scores were treated using the analysis of covariance technique. Analysis of covariance is an extension of analysis of variance which permits adjustment of scores on the dependent variable, in this case achievement test scores, for initial differences in pre-test scores. Assumptions underlying covariance include random assignment of subjects, homogeneity of variance, and normality or near-normality of distribution.

In this experiment, administration of a pre-test was a requirement for participation by all the universities involved. Since pre-test scores were available, analysis of covariance was selected as a statistical technique rather than an analysis of variance based on gain scores.

As Edwards (1960) points out, analysis of variance based on gain scores assumes a common regression coefficient equal to 1.00. Where this assumption is not met, as is usually the case in research, then analysis of covariance provides greater control of experimental error than analysis of variance based on gain. In this instance, the within groups regression coefficient of .5713 clearly indicated the use of covariance. Where its assumptions have been met, covariance is generally accepted as a more powerful statistical technique and one that adds greater precision to analysis of the data.

Inspection of the scores arranged in a distribution table indicated that the assumption of normality of distribution was tenable. Homogeneity of variance within each treatment group was tested using the procedures outlined by Edwards (1958, p. 328). None of the F values calculated in this manner reached the .10 level of significance. Therefore, the assumption of homogeneity of variance was accepted.

In applying covariance the procedures outlined by Garrett (1953, pp. 290-296) were followed. Sums of squares for the total, among group means, and within groups were determined. At this point an analysis of variance of the pre-test and the achievement test scores separately was conducted.

The analysis of variance of the pre-test scores gave an F ratio of 1.81 which, with 2 and 89 degrees of freedom, did not reach significance at the .05 level. This non-significant F value indicated that the random assignment of subjects to the various experimental groups had operated successfully.

For the achievement test scores, the analysis of variance produced an F ratio of 3.28. This F was significant at the .05 level with 2 and 89 degrees of freedom.

The sums of squares for the achievement test (Y) were then adjusted for initial differences in pre-test scores (X) and the F ratio of Y adjusted for X computed. The resulting F ratio of 4.13 proved significant at beyond the .05 level of confidence (2 and 88 of freedom.)

At this point that part of the null hypothesis regarding the effect of the various treatments on achievement could be tentatively rejected. The significant F ratio indicated that at least one treatment mean differed from the others at a level beyond what would ordinarily be expected by chance alone.

Regression coefficients were calculated and the achievement test means adjusted for initial differences using the within groups coefficient of .5713 on the pre-test. The results of this adjustment are summarized in Table 17.

TABLE 15

SUMMARY OF ANALYSIS OF VARIANCE OF PRE-TEST (X) AND ACHIEVEMENT TEST (Y) SCORES

Source of Variation	df.	Sum of Squares (x)	Sum of Squares (y)	Mean Square (x)	Mean Square (y)	Fx	p	Fx	p
Among Means	2	74	194	37.0	97.0	1.81	(N.S.)	3.28	(Sig. .05)
Within Groups	89	1,815	2,636	20.4	29.6				
TOTAL	91	1,889	2,830						

TABLE 16

SUMMARY OF ANALYSIS OF COVARIANCE OF PRE-TEST AND ACHIEVEMENT TEST SCORES

Source of Variation	df.	Sum of Squares (x)	Sum of Squares (y)	Sum of Squares (xy)	Sum of Squares (y Adj.)	Mean Square y (Adj.)	SD y (Adj.)	F y (adj.)	p
Among Means	2	74	194	24	190	95	4.74	4.137	(Sig. .05)
Within Groups	88	1,815	2,636	1,037	2,044	23			
TOTAL	90	1,889	2,830	1,061	2,234				

TABLE 17

PRE-TEST, ACHIEVEMENT TEST, AND ADJUSTED ACHIEVEMENT TEST MEANS FOR THE THREE EXPERIMENTAL GROUPS

Group	N	Pre-Test Mean	Achievement Test Mean	Adjusted Achievement Test Means
Overt	28	13.29	28.82	28.28
Covert	33	11.18	27.03	27.70
Discussion	31	12.74	25.19	24.92
General Means		12.35	26.96	26.94

To determine which mean or means differed significantly Tukey's test for a significant gap was applied. (Edwards, 1958, p. 331) Some difficulties were encountered since the N in each group differed slightly. In this situation Winer (1962, p. 101) suggests that a harmonic n can be used provided the difference in the size of the N's is not too great. A harmonic N of 30.54, calculated using the formula in Winer, (1962) was used in the denominator of the Tukey formula to determine the size of gap between means required for significance at the .05 level. This procedure indicated the required gap as 2.44 or more. Inspection of the differences between means indicated that both the overt and covert adjusted group means differed significantly from the discussion group adjusted mean at the level selected for significance (.05.) There was no significant difference between the overt and covert group means.

TABLE 18

SUMMARY OF TUKEY TEST FOR A SIGNIFICANT GAP AMONG MEANS

Comparison	Difference in Adjusted Means	Difference For Significance (.05)	P
Overt vs. Covert	.58	2.44	(N.S.)
Overt vs. Discussion	3.36	2.44	(Sig. .05)
Covert vs. Discussion	2.78	2.44	(Sig. .05)

Thus the null hypothesis regarding achievement can be rejected at the .05 level of confidence since both the overt response group and the covert response group performance proved significantly superior on the achievement test measure.

2. Instructional approach and retention - Pre-test and retention test scores were analyzed using covariance procedures similar to those described previously. Due to absences at the time of retention testing, the size of the overt, covert, and discussion groups was reduced from 28, 33, and 31 to 27, 31, and 28 respectively. As a check, another analysis of covariance of the achievement test results was run with the scores of the retention test absentees withdrawn from the sample. Elimination of the scores of the 6 subjects absent did not affect the findings of the original analysis of covariance of achievement. Conversely, it does not follow that the absence of these 6 subjects had no effect on the retention test means.

In the covariance analysis Garrett's (1953) procedures were again followed. The preliminary analysis of variance of the pre-test (X) and retention test (Y) scores gave a non-significant F ratio of 2.12 for the "X" scores and an F ratio of 3.33, significant at the .05 level, for the "Y" scores.

Adjusting the sums of squares for the retention test (Y) in terms of initial differences on the pre-test (X) gave an F ratio of 2.25 which did not prove significant at the .05 level with 2 and 84 degrees of freedom. Thus the null hypothesis regarding retention was not rejected since retention test scores among the treatment groups did not differ significantly when adjusted for differences on pre-test scores.

At this point it should be noted that three of the four university groups involved in this experiment received regular instruction in school law during the interim between achievement and retention testing. It can be assumed that the instructional procedures used would more closely resemble the case study-discussion approach than either of the programming approaches. In a sense then the post achievement test instruction may have been more reinforcing to the discussion group than to either of the programming groups. Although the subjects were randomly assigned within schools, the interim instruction may have served to mask significant differences present at the conclusion of the experiment.

TABLE 19

SUMMARY OF ANALYSIS OF VARIANCE OF PRE-TEST AND RETENTION TEST SCORES

Source of Variation	df.	Sum of Squares (x)	Sum of Squares (y)	Mean Square (x)	Mean Square (y)	Fx	p	Fy	p
Among Means	2	88	185	44	92.50	2.12	(N.S.)	3.33	(Sig. .05)
Within Groups	83	1,723	2,305	20.76	27.77				
TOTAL	85	1,811	2,490						

TABLE 20

SUMMARY OF ANALYSIS OF COVARIANCE OF PRE-TEST (X) AND RETENTION TEST (Y) SCORES

Source of Variation	df.	Sum of Squares (x)	Sum of Squares (y)	Sum of Squares (xy)	Sum of Squares (y Adj.)	Mean Square (y Adj.)	SD (y Adj.)	F (y Adj.)	P
Among Means	2	88	185	100	103	51.50	4.79	2.25	(N.S.)
Within Groups	82	1,723	2,305	856	1,880	22.93			
TOTAL	84	1,811	2,490	956	1,983				

TABLE 21

SUMMARY OF PRE-TEST MEANS, ACHIEVEMENT TEST MEANS (UNADJUSTED) AND RETENTION TEST MEANS BY TREATMENT GROUP

Group	Pre-Test Mean	Achievement Test Mean	Retention Test Mean
Overt	13.29	28.82	25.04
Covert	11.18	27.03	21.97
Discussion	12.74	25.19	21.79

3. Attitude scale measures and type of experimental treatment - Usable opinion instruments were returned by 92 subjects of 103 present when they were distributed. Among these were several late enrollees at universities "B" and "C" who had not been present for the pre-test or assignment to experimental groups. They did participate in the principal experiment, however, and it was felt that their opinions were a reflection of a valid exposure to the treatments under study.

The total score for each subject on the Likert scale was determined and these scores were tested for significant differences in group means using simple analysis of variance procedures.

Although the discussion group mean exceeded that of either the overt or covert group, the analysis of variance indicated no difference among the groups in their opinions of the treatments to which they were exposed.

In failing to reject the "no difference" hypothesis in this regard it cannot absolutely be assumed that no real differences in opinion existed among the various groups. It can be stated that using a Likert type instrument of high reliability, no significant differences in opinion of experimental treatment could be detected among the various groups.

TABLE 22

RANGE OF SCORES AND MEAN SCORES ON THE LIKERT SCALE BY EXPERIMENTAL GROUPS

Group	N	Range of Scores	Mean Score
Overt	30	33 - 77	59.53
Covert	31	34 - 78	59.93
Discussion	31	41 - 80	62.64

TABLE 23

SUMMARY OF ANALYSIS OF VARIANCE OF LIKERT SCALE

Source of Variation	d.f.	Sums of Squares	Mean Square	F	p
Among Means	2	138	69	.0176	(N.S.)
Within Groups	89	348,205	3,912		
TOTAL	91	348,343			

4. Relationship of attitude to achievement and retention - To test for the significance of relationships between opinion of an instructional method and achievement in that method a series of product moment correlations were calculated. The significance of the correlation was determined by referring directly to Table VI in Edwards (1958) with N-2 degrees of freedom. Similar procedures were followed using retention test and opinion scores. Results of these correlations are reported in Table 24.

TABLE 24

PRODUCT MOMENT CORRELATIONS BETWEEN ATTITUDE AND ACHIEVEMENT AND ATTITUDE AND RETENTION FOR THE THREE TREATMENT GROUPS

Variable	Overt Group			Covert Group			Discussion Group		
	n	r	p	n	r	p	n	r	p
Attitude and Achievement	26	-.054	(N.S.)	29	.107	(N.S.)	27	.070	(N.S.)
Attitude and Retention	26	-.118	(N.S.)	27	.207	(N.S.)	26	.105	(N.S.)

On the basis of these data the null hypothesis regarding the relationship of opinion to achievement and retention was not rejected. The null hypothesis holds for all treatment groups. The slight negative correlation between opinion, achievement, and retention for the overt response group should be noted. This may reflect a tendency on the part of more able students to feel deterred or held back by the necessity of responding overtly to each item. On the other hand, factors within the experiment itself may have produced this result. In follow up interviews, members of this group commented on "how hard they had worked," "how fatigued they were," etc. Their reaction to the attitude instrument given shortly after the experiment may be mainly a reflection of these feelings.

Conversely, it is interesting to speculate on the opinion of low performance students in the overt group. The non-significant negative correlation of opinion and achievement hints that these students may have been more favorably impressed by the step by step thoroughness of the program and the high level of reinforcement received. No further conclusions can be drawn since statistically the relationship represents a chance occurrence.

5. Summary of personal interviews - A total of 22 subjects or 24% of the sample was interviewed subsequent to the completion of the experiment. All subjects selected for interview scored either at the 10th percentile or below or at the 90th percentile or above on either the achievement measure or the opinion scale. Only one subject, a low achieving member of the covert response group, indicated an unwillingness to make arrangements for an interview.

Those selected for interview included 6 each from the high achievement group, the high opinion group, and the low achievement group and 5 subjects who expressed very low opinions of the experimental method to which they were exposed.

a. Areas of disagreement among programming groups

1.) Importance of the time factor - About 2/3 of those in the low achievement and low opinion groups expressed concern or dissatisfaction with the amount of time allotted for the experiment. This factor was mentioned as important by only 3 members of the high achieving and high opinion groups. Differences in reading speed may be an underlying factor here.

2.) Retention - Those who ranked high in achievement or in their opinion of the program generally felt that their retention of the programmed material was "better than" the rest of the course.

One high opinion covert response subject stated:

> "In the final exam I remembered the questions taught by programmed instruction; they seemed to stick better."

Low opinion subjects and low achievers generally stated their retention was "the same as" or "lower than" the rest of the course. Two commented that they thought their retention was better in discussion groups than in the programming group.

3.) Program content - High achievers and members of the high opinion group generally agreed that the content of the material was satisfactory.

One high achiever, disagreeing to some extent, commented:

> "It was very well organized but a little too babyish for me with too much repetition... Too logical for me... In a lecture they leave gaps so that you can think for yourself. I like to think for myself."

Others in the high groups commented generally on how well the material was organized.

The low opinion group was generally critical of the content. One stated:

> "It was not organized as well as it might have been."

Another low opinion subject, a law school graduate who had performed quite well on the achievement measure, felt that many of the answers were misleading. He didn't like being "boxed in" on the answers and complained that many answers were "black or white when they should have been a shade of grey." He didn't believe that law generally was a subject amenable to programmed instruction.

Another low achiever would have preferred more current and up-to-date material rather than, as he put it, the "historical" background of Supreme Court decisions.

4.) Number and length of case studies - The majority of the high achievement and both high and low opinion groups favored more case studies.

All low achievers indicated they would prefer fewer case studies. One commented, "I'm not sure the case studies were any good."

The low opinion subject with law school training would have preferred more case studies but thought that the selection of cases was poor and the editing biased.

5.) Manner of use - The high achievement and high opinion groups had little objection to occasional or even frequent use of the programmed material during class time. Low achievers and those with low opinions of the material definitely preferred use at home or out of class.

One high achiever commented that he had no objection to use in class and added:

> "If a lecture is well planned it is as good as a program."

One member of the low opinion group stated:

> "Use it for homework. If I am paying for a course why should I sit and do it there?" (in class)

b. Areas of agreement among programming groups.

1.) Satisfaction with group assignment - Very few of those interviewed would have preferred to have been assigned to another experimental group. Several stated that they believed the overt group had an advantage in that they read and wrote the responses.

One low opinion covert group subject stated: "We had the speed but I learn better when I write." An overt response low achiever who classified himself as a slow reader felt that he "would have had more of a chance with respect to time" had he been assigned to the covert response group.

2.) Previous experience with programming - Only three of those interviewed had previous experience with programmed material; one had written part of a program in distributive education, another had taken graduate work with a programmer, and a third has supervised experimental use of programmed material in algebra in his high school.

3.) Program format - All subjects interviewed agreed that this was satisfactory. One commented that the tabs were easier to handle than a sliding mask.

4.) Attitudes towards programmed instruction before and after the experiment - No great difference among the groups was evident here. High achievers were either neutral or favorable in their attitude before the experiment and favorable in their attitudes afterwards. Even low opinion subjects admitted programmed instruction had limited usefulness, although two of those interviewed mentioned boredom and lack of motivation as detrimental factors.

5.) Importance of reading speed and comprehension - There was general agreement as to the importance of these factors, particularly comprehension, in working successfully with this material.

One low achiever commented, "Very important I know that I am a slow reader and time pressed."

6.) Place in graduate work - There was general agreement that there was a place for programming in graduate classes. Opinions differed as to how best to use the material.

7.) Reasons for graduate study - There were more doctoral candidates among the high achieving and high opinion groups. Generally members of the low achievement and low opinion groups were older, two were in their fifties, and taking graduate work for certification or as a "refresher."

8.) Undergraduate preparation - Members of each group came from a variety of undergraduate institutions; universities, liberal arts schools, and state colleges. No trends associated with achievement or opinion were noticeable here.

c. Features of the experiment most liked and disliked - High achievers frequently mentioned their satisfaction with the amount of material learned. Typical comments included:

> "... I felt when I was finished that I had really learned what was presented and that I would retain it.
>
> I liked most the fact that I learned something."

One high achiever commented that he disliked the fact that the program forced him to "work harder." Others disliked the time factors under which the experiment was conducted.

One high opinion subject who had not attended graduate school for 15 years commented that "Practicing this (the program) was worth the entire course to me."

Two low achievers expressed the view that more notice of the experiment would have helped them develop a better "mind-set" to the whole affair.

Boredom and distaste for use in class were mentioned as dislikes by members of the low opinion group. One member of this group liked having the opportunity to work with programmed material although he disagreed with many of the responses.

d. Summary of interviews with discussion group members - A majority of the interviewees in the discussion group felt that the time factor was disturbing. They would have preferred fewer case studies, perhaps treated in greater depth.

Those who had achieved well or whose opinions of the method were above average generally rated the discussions high in quality. Low achievers thought the discussions were only average or below average in quality.

Comments regarding the discussion guide were very favorable. "Terrific" and "added very much" were typical of the comments. Apparently preparation of discussion guides or outlines alone can do much to add to learning in and student satisfaction with discussion sessions.

The discussion guide was frequently mentioned as a most liked item. Some comments:

> "I can envision a whole course being taught with discussion guides and case studies.
>
> I liked the discussion guide. It was a good crutch."

No member of the programming groups endorsed programming to the extent of suggesting it be used for an entire course.

Discussion group interviewees generally rated the discussion guide as most helpful and the lack of opportunity for outside preparation as the least liked part of the experiment.

e. Some conclusions from the interviews - Apparently a combination of case studies and well designed discussion guides combines the properties of efficiency in learning with a high degree of acceptance by students. It requires the services of a person knowledgable in the subject field as discussion leader, whereas the programmed material is a self-contained instructional unit. Choice of methods might well depend on the availability of instructors and the emphasis placed on the importance of group processes in graduate work. Combining the opinions of students interviewed, the most satisfactory method might include a combination of programmed material and case studies for use outside of class together with well planned discussions under the leadership of a subject expert during class time.

C. Secondary Data

Findings not directly connected with the elements of the problem are reported in this section. Some of this data shed further light on the findings previously reported. Other information indicates possible directions for future research.

1. Teaching experience, achievement and retention - To determine the nature, extent, and possible significance of relationships between teaching experience, achievement and retention a series of rank order correlations was calculated for each experimental group. Subjects within each group were ranked from high to low on each variable and the correlations calculated following procedures outlined by Edwards (1958, pp. 427-428). In each case the correction factor for tied ranks was applied. Significance was determined by referring directly to Table VI in Edwards (1958) with N-2 degrees of freedom.

These correlations are reported in Table 25.

No relationship between experience and achievement or retention reaching significance at the .05 level could be determined for any group. Among the overt group subjects, there was a negative correlation between achievement and years of teaching experience, significant at the .10 level. Physical factors may have entered into this finding. Younger subjects generally would have better reflexes and greater physical stamina. Therefore, they should be more able to work with this material over the long continuous block of time involved in this experiment. On the other hand, this negative correlation could also reflect opposition on the part of more mature students to a continuous schedule requiring overt responses. Both of these possibilities can be offered in an attempt to explain this negative tendency.

TABLE 25

RANK ORDER CORRELATIONS BETWEEN YEARS OF TEACHING EXPERIENCE; ACHIEVEMENT AND RETENTION

Variables	Overt Group			Covert Group			Discussion Group		
	n	r'	p	n	r'	p	n	r'	p
Experience and Achievement	28	-.369	(Sig. .10)	33	.086	(N.S.)	30	.098	(N.S.)
Experience and ~~Achievement~~ RETENTION	26	-.253	(N.S.)	27	.042	(N.S.)	26	.010	(N.S.)

2. Graduate credits earned, achievement and retention - No significant relationship was evident between the number of graduate credits earned and achievement or retention. Programmed material of the type developed for this experiment can evidently be used successfully with populations which are quite heterogeneous in terms of graduate level preparation. In number of graduate credits earned, subjects in this experiment ranged from 0 to over 100 credits.

TABLE 26

RANK ORDER CORRELATIONS BETWEEN GRADUATE CREDITS EARNED, ACHIEVEMENT AND RETENTION

	Overt Group			Covert Group			Discussion Group		
Variable	n	r'	p	n	r'	p	n	r'	p
Credits Earned and Achievement	28	.171	(N.S.)	31	-.246	(N.S.)	30	.206	(N.S.)
Credits Earned and Retention	26	.111	(N.S.)	27	-.042	(N.S.)	26	.169	(N.S.)

3. Experimental treatment and completion rate - In planning the experiment, consideration was given to collecting data regarding the rate at which overt and covert response subjects completed the programmed material and case studies. The proctors had been instructed to announce the time at 50 minute intervals and to ask students to record this time in the programmed material or case studies. However, in one instance these instructions had not been given at the requested time and several subjects either failed to record the time or recorded it illegibly.

The problem of estimating completion rates was further complicated by the nature of the materials. Case study materials were integrated with the programmed material in such a way that the number and length of the case studies varied from unit to unit. Thus a time estimate based on frames completed alone would have been quite misleading. Further, it can be assumed that most of the subjects, particularly in the overt response group, completed the case studies at a different and faster rate than the program.

To adapt the available time data to statistical treatment, a sample set of programmed materials and case studies was numbered in order from 1 to 79 with each number representing approximately one page of either programmed material or case studies. The proper number was then divided by the number of minutes represented in the time data reported by each subject. The resulting ratio, referred to here as a time-completion ratio, was an index of the speed at which each subject proceeded through the materials. A student with a time completion ratio of .40 for example completed approximately 2/5 or 40% of a page per minute. Ratios varied from .77 for the fastest subject to .21 for the slowest student. The use of these ratios made possible certain comparisons in the rates at which subjects in the two programming groups completed the material.

An inspection of the materials used by the subjects showed that only a few had recorded the time at the 150 minute mark or at the conclusion of the program. However, a large majority of the subjects had properly indicated the time at the 50 and 100 minute marks (65 and 115 minutes at one university.) Time completion ratios were calculated for both the first and second times recorded. Since each subjects' time completion ratios were quite similar at the intervals checked it was decided to base any further statistical treatment on the ratios calculated from the times recorded at or near the 100 minute mark. The mean time completion ratio for the overt response group was .36 and for the covert response group .45.

A comparison of the overt response group and the covert response group time completion ratios was made using the Kruskal-Wallis Test as outlined in Edwards. (1958, p. 425) Application of this test yielded a H value of 9.217. In the χ^2 table with 1 degree of freedom this H value was significant at the .001 level. We can conclude, therefore, that the overt and covert response groups differed significantly in their rates of progress through the programmed material. This supports findings in other studies where covert response has been found to be more efficient in terms of time invested. It also raises questions as to why this efficiency is not reflected either in higher achievement or better retention. The covert response group in this situation had additional time for study and review. The proctors reported that the entire amount of time available was spent with the programmed material. Yet apparently covert response was no more efficient than overt response on the criterion measures when the time factor was held constant for both. Several of the overt response subjects in personal interviews "felt" that writing the answers assisted them in learning since it involved them more deeply in the material. Where time is not a major factor, overt response may have motivational or reinforcing values not measured by any of the criterion instruments used in this experiment.

TABLE 27

RANGE OF TIME-COMPLETION RATIOS AND MEAN TIME-COMPLETION RATIOS FOR OVERT AND COVERT RESPONSE GROUPS

Group	Range of Time Completion Ratios		Mean Time Completion Ratios
	Low	High	
Overt	.32	.52	.36
Covert	.21	.77	.45

4. Further analysis of the program - The rank order correlation procedure was also used to test the relationship between total program error rate and performance on the achievement test. Based on the data secured from the overt response group, the correlation between the total program error rate and achievement was .213. This correlation did not prove significant at the .05 level.

The program items classified as criterion items were checked separately and rankings based only on errors on criterion items were established. The rank order correlation between criterion item error rate and achievement rank was in this case determined to be .285. This correlation, although it did show a greater degree of relationship, did not reach the level of significance required at the .05 level.

Since the achievement test was administered immediately after work with the programmed material was completed, there is a possibility that students missing criterion items "learned from mistakes immediately corrected" in the program and then transferred this learning directly to the achievement test. This is not an argument against controlled program error rate, particularly on criterion items. Careful development of the program and elimination of confusing or misleading items is most essential. However, even with these conditions, students still learn from incorrect responses, provided immediate knowledge of results is provided.

Finally, it should be noted that the program used was classified as a generally low error rate program. Individual student criterion item error rates ranged from .000 to .157 for the student with the greatest number of errors. The relationship of correct response on criterion items to achievement may be more evident in programs with a greater difficulty level.

5. Student opinions regarding the case studies used in the experiment.

a. Opinions as to interest - All of the subjects responding to the questionnaire, regardless of treatment group, rated the case studies as "extremely interesting" or "rather interesting."

b. Opinions as to value in the learning situation - Of those responding 97% from the overt response group indicated that the case studies "were of some value" or "added a great deal to the learning situation." All from the other treatment groups responded similarly. Only 3% of the overt response group indicated indecision in regard to this factor and no student from any treatment group indicated a negative opinion.

c. Opinions regarding style and manner of expression - Of the overt response group 87%, of the covert response group 88%, and of the discussion group 80% rated the case studies as "above average" or "very well written." Of the three groups 13%, 6%, and 13.3% respectively checked "undecided" in responding to this item.

d. Opinions regarding the length - A large majority of the respondents, regardless of treatment group, indicated that the case studies "were just about the right length." Of the overt response group 13.3% as compared with 10% of the covert response group and only 3.3% of the discussion group rated the case studies as "too long and detailed."

e. Additional student comments regarding the case studies - Additional comments were offered by 13 members of the overt response group (43%), 10 members of the covert group (32%), and 14 members of the discussion group (47%).

Of the overt responders, five indicated lack of time as a problem area. Time was not a major factor in covert group comments. Several commented on typographical errors in the case studies.

The most extensive comments were made by members of the discussion group where the case studies were the major learning aid. Their comments were largely very favorable or quite favorable in nature.

6. Student opinions regarding the programmed material used in the experiment - Only members of the overt and covert groups responded to this section of the questionnaire. Replies were received from a total of 61 students who had worked with the program. (Overt - 30, Covert - 31)

a. Interest level of the program - Among the overt group the program was rated as an "extremely interesting one to read" by 23% and as "rather interesting reading" by 60%. Similar responses were checked by 29% and 58% of the covert group respondees. Ten percent of the overt group and 6.5% of the covert group indicated that the program was "somewhat boring."

b. Size of step - Twenty-three percent of the overt group and 16% of the covert group indicated that the frames or steps in the program were "large enough to be challenging." "Just about right" in response to this item was checked by 57% of the overt group and 74% of the covert group.

c. Program difficulty - Seventy-seven percent of the subjects responding in each group rated the program as "just about right" in difficulty. Seven percent of the overt group subjects and 13% of the covert group subjects rated the program as "rather difficult."

d. Physical arrangement - Of the overt group 86.6% and of the covert group 77% indicated that the format of the program "worked very efficiently" or "seemed O.K." Ten percent of the overt responders and 23% of the covert responders indicated that the "format caused some problems."

e. Value in the learning situation - According to 53.3% of the overt group and 55% of the covert group the programmed material "added a great deal to the learning situation." Thirty percent and 36% of the groups respectively indicated that the material was of some value. Of the overt responders 16.6% and of the covert responders 6% felt that the material was either of questionable value or detracted from the learning situation.

f. The program and class discussion - Eighty-one percent of the overt responders and 77% of the covert responders missed the give and take of class discussion to a greater or lesser degree. Only 4 subjects in each group (13%) indicated they did not miss class discussion.

g. The program and fatigue - Among those students recording their responses, over half (63.3%) "began to feel fatigued" at the 100 minute point or earlier. A similar opinion was expressed by 45% of the covert responders. Twice as many covert responders (26%) as overt responders (13.3%) had no feeling of fatigue during the entire experiment.

h Additional student comments regarding the programmed material - Additional comments were given by 19 members (63%) of the overt response group and 21 members (68%) of the covert response group.

Among the overt group the factors of time and fatigue again received major attention. Other comments indicated the need for variation in teaching technique, with several suggesting combinations of programmed material with discussions.

Three of the covert response group mentioned the fatigue factor in their comments. Several recommended shorter units or shorter periods of use, combined with other teaching approaches. Three respondees indicated they would have preferred a written response approach.

One comment indicated that the reliability of the completion times listed by the covert group might be open to question.

> "I started part A too fast, was finished in less than half the time, deliberately slowed down to meet time schedule."

7. Student opinions regarding potential uses of the programmed material.

a. Suitability for use on the graduate level - Of the overt response group 86% indicated that the program should be used by "most" or "all" graduate students. Only two subjects in this group (7%) absolutely opposed its use in graduate work. Opinions of the covert group were more diversified. While 58% of this group supported its use by "some" or "all" graduate students, 16% were "undecided" and another 19% indicated that the program seemed "suitable for only a few graduate students."

b. Manner of use - Of the overt group 77% and of the covert group 69% favored "occasional use during class time." Continual use during class time was favored by 7% and 10% of the treatment groups respectively. Only one subject in each group (3%) felt that class time should never be used for this material.

c. The program vs. the conventional text as a study device - Of the overt response group 77% signified that the program was "superior to" or "better than" a conventional text. A similar opinion was expressed by 61% of the covert response group. The program was rated as "inferior to a conventional text" by 7% of the overt group and 3% of the covert group. Of the covert group 13% rated the program and a conventional text as "about the same."

d. Recommendations for further development - Among this covert group respondees, 34% recommended development of the programmed material for other graduate courses; an opinion shared by 17% of the overt response subjects. An additional 73% of the overt group and 40% of the covert group recommended development of such material for both undergraduate and graduate work. Only one subject in each group supported restricting programmed material to the elementary or secondary level.

e. Additional student comments - Comments regarding possible uses of the material were made by 15 students (50%) of the overt response group and a similar number (48%) from the covert response group. Several respondents from both groups indicated that, while they were impressed with programming as a medium for factual material, they had reservations about its usefulness in concept-building, problem solving or "creative thinking" courses.

Comments from the covert group regarding uses of this material extended into other areas including who should control program development and the role of the instructor in learning. Many saw programming as acceptable if used in combination with other methods.

8. Student opinions regarding the discussion approach - Only students who participated in the discussion sessions responded to this section of the questionnaire. Twenty-six discussion group subjects returned usable questionnaires.

a. Organization of the discussion guides - All subjects rated the discussion guides as "about average," "well organized," or "very well organized."

b. Value of the guides in the learning situation - Of those responding 61% rated the guides as "very helpful" and an additional 35% indicated they "were of some help."

c. Quality of the discussions - The discussions were rated as "about average" by 50% of those who replied. Another 35% indicated the discussions were "above average in stimulating thought." Of the group 15% thought the discussions were "very stimulating."

d. Group participation in the discussions - Participation was rated as "well above average" by 12% of those responding and as "about average" by 69%. One subject (4%) thought that participation was "very poor" and 15% indicated that participation was "below what would be normally expected."

e. Additional comments - Thirteen members of the discussion group added comments, some of which were rather extensive. Several mentioned that the discussions would have been more effective had they prepared in advance. Lack of time was again mentioned as a factor by a few respondents. Several subjects expressed satisfaction with the discussion procedures.

9. Analysis of a typical case study discussion session - In working with the discussion groups the experimenter attempted to equate the discussion experience as much as possible with that of the programming groups. Content was limited by a specially prepared discussion guide to the same material covered in the program. The material was covered in the same order; the same case studies were used and, as far as possible, time for each discussion unit was equated with the time spent in similar units of the program. The time data collected from subjects in the development and testing of the programmed material and case studies proved helpful in establishing these standards. At each school the programming and discussion groups spent the same amount of time overall in the experimental situation.

In the discussions themselves the experimenter consciously attempted to emulate the reinforcing properties of the program. Questions were asked in such a manner as to stimulate as much student response as possible. Verbal reinforcements such as "Very Good," "Fine," "That's correct" were used whenever possible and negative type comments were minimized.

Despite the rigorous attempts to equate the groups as to content material covered, discussion group members on several occasions introduced material not contained in either the discussion guide or the program. For example, at the time of the experiment, the admission of James Meredith to the University of Mississippi was an item of current interest. The legal questions involved in Meredith's case were introduced by discussion group members at all the schools involved in the experiment.

From the viewpoint of research control, the introduction of extraneous material may have caused some difficulty since time spent on these topics would not contribute directly to learning as measured by the achievement instrument. However, the flexibility of the case study discussion approach permitted a form of "branching" to include current topics. Similar flexibility with the programming groups would require the intervention of an instructor or continuous revision of the program to keep it up to date.

In dealing with extraneous material the experimenter attempted to get the groups back to the content of discussion guide at the first opportune moment. Discussion of extraneous topics was not shut off immediately since this might have interfered with the willingness of subjects to respond to other material.

Table 28 summarizes the types of activities carried on in one case study-discussion session. To construct this table the experimenter listened to a tape recording of the session and categorized the activities under the various headings listed.

It should be noted that the discussion leader gave only 39 reinforcements in the form of verbal praise or comments while the program provided for over 400 reinforcements in the form of knowledge of results. However, in the discussion group, correct student responses (69) and many of the student comments (123) could be classified as reinforcing in nature. Further, reinforcements given in social activities such as discussion may be more effective than individual knowledge of results since they are given in the presence of the entire group. Commendations or correct responses given publicly may be much more strongly reinforcing than if given only privately to each subject.

The number of student questions and comments indicated a high level of participation on the part of the subjects involved. This may reflect both the organization and preparation for the discussion session and an understanding by the subjects of their involvement in an experimental situation.

TABLE 28

SUMMARY TABLE OF TYPES OF DISCUSSION ACTIVITIES AND FREQUENCY OF EACH IN A TYPICAL DISCUSSION SESSION

(UNIVERSITY A)

TYPE OF ACTIVITY

Discussion Guide Unit	Questions by Discussion Leader	Student Responses		Reinforcing Type Comments By Discussion Leader	Student Comments	Student Questions	Cases Read
		Correct	Incorrect				
I	34	29	4	19	18	5	0
II	17	13	2	6	14	5	2
III	5	3	5	1	--	12	1
IV	18	13	3	6	70	10	3
V	14	11	2	7	21	--	1
TOTAL	88	69	16	39	123	32	

V. FINDINGS, CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY

A. Findings

Findings regarding the elements of the problem are reported in the section immediately following. Additional findings classified as secondary to the major purposes of the research are reported in section two.

1. Findings regarding the elements of the problem

a. The null hypothesis regarding achievement must be rejected. Both the case study-programmed-overt response group and the case study-programmed-covert response group performed better on the achievement measure than did the case study-discussion group. This superiority was maintained with the achievement test scores adjusted for initial differences among the groups as measured by the pre-test. The achievement of the two programming groups did not differ.

b. In the retention test the best performance was recorded by the programmed-case study-overt response group followed by the programmed-case study-covert response group and the case study-discussion group. However, with the retention test scores adjusted for pre-test scores, the groups did not differ in retention.

Results regarding retention must be interpreted cautiously since at three schools participating in the experiment regular instruction in school law continued during the interim between achievement and retention testing.

c. As measured by a Likert type attitude instrument no difference could be found among the various groups in attitude toward the experimental method to which they were exposed. In other words, the opinion of the programming groups regarding their methods of instruction was as favorable as that of the discussion group toward its method of instruction.

d. Attitude toward an experimental method was not significantly correlated with either achievement or retention of material taught via the method.

e. Interviews were arranged with 22 subjects selected from those scoring at the 90th percentile and above on either the achievement measure or the opinion scale and at the 10th percentile or below on these measures. Interviewees generally were satisfied with their experimental group assignment. Very few had any previous experience with programmed instruction other than professional reading. All subjects interviewed from the programming groups agreed that the program format was satisfactory. This exposure to programmed instruction, either with overt or covert response, did not affect the attitudes of interviewees toward programming unfavorably. High achievers and high opinion subjects expressed their attitude toward programming as favorable or highly favorable as a result of this experiment. Low opinion subjects or low achievers expressed neutral or only slightly favorable attitudes. All subjects interviewed from the programming groups agreed that reading speed and comprehension were significant factors in working successfully with this material. There was general agreement that there was a place for programmed material in graduate work, although opinions as to manner of use differed. Reasons for graduate study in general and studying school law in particular were not associated with either high or low opinion or achievement. High or low opinion or achievement was not related to the type of institution where the undergraduate preparation was received.

f. The high opinion-high achievement groups differed from the low opinion-low achievement groups in their evaluation of the importance of the time facto- in the experiment. Generally, low achievement or low opinion subjects felt pressed for time and expressed concern about this factor.

g. Opinions regarding retention also differed among the high and low groups, with the high groups generally stating they retained more from the program and the low groups stating they retained less.

h. Low achieving and low opinion interviewees were more critical of the content of the program than those from the high achieving - high opinion groups.

i. High achievers and those expressing highly favorable and unfavorable opinions preferred more case studies, while those from the low achievement groups would have preferred fewer case studies. This reaction on the part of low achievers may be a reflection of certain inadequacies in reading speed or comprehension.

j. Interviewees generally would have preferred that the case studies be shorter; an opinion which differed from that expressed in the written opinion instrument.

k. High scoring subjects on the achievement or attitude scale had little objection to occasional or even frequent use of programmed material during class time. Low scorers on either variable generally preferred use only outside of class hours.

l. Interviewees from the discussion groups were very favorably impressed with the discussion guides. Generally they would have preferred more time for the discussion session and some opportunity for advance preparation outside of class. Low opinion interviewees from the discussion group mentioned such factors as the experimental nature of the whole situation and poor physical arrangements for the discussion group as detrimental factors.

2. Secondary Findings

a. A negative correlation between achievement and years of teaching experience, significant at the .10 level, was found for the case study-programmed-overt response group. This negative correlation held for retention but not at a significant level. For the covert response group and the discussion group, years of teaching experience was not correlated with either achievement or retention.

b. Neither achievement or retention by any group was correlated with number of graduate credits earned.

c. The case study-programmed-covert response group completed the material at a faster rate than the case study-programmed-overt response group. The difference here was significant at the .001 level.

d. The correlation between either total program error rate or criterion item error rate and achievement did not reach significance. This finding is based on an analysis of the data from the overt response group.

e. The majority of subjects, regardless of experimental group assignment, expressed opinions that the case studies were interesting, of value in the learning situation, well written and about the right length. Follow-up interviews with 20% of the subjects indicated that shorter case studies would have been preferred, at least by those whose opinions or achievement represented extremes.

f. A majority of the overt and covert group responders expressed favorable opinions regarding the characteristics and format of the programmed material.

g. Majorities in both groups missed the give and take of class discussion, indicating the importance placed by these subjects on social contact and exchange in learning.

h. The fatigue factor became of concern at about the 100 minute mark in the use of the programmed material. Fatigue was a problem for a larger proportion of overt responders than covert responders.

i. A majority of both the overt and covert group approved of the use of this type of programmed material by graduate students and rated it as superior to the conventional text.

j. In written responses both overt and covert groups favored occasional use of programmed material during class time.

k. A majority of the discussion group members approved of the discussion guides and rated them as very helpful in the learning situation. A majority rated both the quality of the discussions and participation in the discussions as average or above.

B. Conclusions

A combination of programmed material and case studies seems quite suitable for limited use as a graduate-level teaching technique. Achievement and retention results, and the highly favorable opinions expressed by the mature graduate students who participated, all indicate the contribution such materials might make in graduate level courses. The case study programming approach should also prove suitable for certain professional preparation in other areas such as medicine, law, and social work.

In analyzing the possibilities of using programmed material and case studies, it should be emphasized that superior achievement and equivalent retention when compared with the case study discussion approach was produced without involving the time of an instructor in any way. Clearly there is at least an economic argument here for the use of the program in that it would permit the more effective use of instructional time of graduate professors.

In developing programs of graduate study, consideration might well be given to completion of selected programmed sequences as a prerequisite for admission to certain courses. Admission to the graduate class in school law, for example, might require the completion of certain units of case study-programmed material and a satisfactory score on an achievement test based on this material. This procedure would encourage the development of habits of independent study and individual responsibility for subject content mastery among graduate students and would assure a certain level of competence and understanding from which the instructor might proceed.

The findings regarding the case study discussion approach are quite interesting. Although this group did not perform as well on the achievement measure, their performance on the retention measure was quite satisfactory. In terms of number of points lost from achievement test mean (unadjusted) to retention test mean, the discussion group lost 3.40 points as compared with 3.78 points for the overt response group and 5.06 points for the covert response group.

The superiority of the programming groups in achievement may be ascribed to the fact that the achievement measure was constructed directly from criterion items in the program, and that the program was more effective than the discussion leader in "teaching" these criterion items. However, the discussion approach permitted flexibility and introduction of topics and concepts not taught in the program or measured by the achievement instrument. This is not necessarily an argument for discussion as it is typically carried out in school situations. It should be emphasized again that the case study-discussion groups used discussion guides prepared in advance and that the discussion leader consciously injected as much verbal reinforcement as possible into the discussions.

Some might offer a Gestaltist type explanation involving better long term integration of all the concepts involved in the topic by the discussion group to explain the "no difference" retention findings. This theory, however, would not explain the superior performance in achievement of the programming groups. The long-term performance of the discussion group may also reflect application of teaching techniques derived from the S-R theory in the discussion sessions.

Regardless of the underlying theory, well-planned discussion carried out with attention to research findings in learning can be both an efficient and satisfying instructional approach. In terms of long-term, continuous, in-class use, it would be more satisfactory as far as students are concerned than either programming approach.

In terms of learning theories, the fact that overt response did not produce significant differences over covert response on the achievement and retention measures tends to support the so-called S-S theory advocated by Tolman and others. However, S-R theorists could point to the slightly superior performance of the overt response group and, more importantly, to the lower loss in retention by both the overt and discussion groups, in support of their theory of the importance of overt behavior in learning. A high level of overt activity was required in both the programmed case study overt response group and the case study discussion group as contrasted with the case study programmed covert response group. Interviews with participants in the experiments indicated that many if not a majority of students believe reading combined with writing to be a superior method of studying or learning. "I learn better with a pencil in my hand," was typical of these comments.

Where time is not a pressing factor, overt response is at least more satisfying to most learners even though its effects in learning may not be detected when compared with covert response. Student satisfaction and security with the subject matter might be sufficient justification for requiring overt response in this situation.

In this research there were indications that more capable students felt constricted by the necessity of overt response. It should be noted in this regard that the program used was of the low error rate non-branching type. Some have suggested that the suitability of overt or covert response may depend on the level of the program error rate, with overt response being more necessary with high error rate or more difficult programs. Suitability and acceptance of response mode may also be related to individual subject's reading skills. Fast readers with good comprehension may feel held back by the necessity of overt response; a feeling not shared by those of only average reading ability.

The failure to find any significant correlation between opinion of an experimental method and achievement or retention may be a reflection of the relatively short exposure of subjects to each experimental method. A study conducted over a more extended period of time might show significant correlations for these variables. The findings in this study are quite similar to those of Eigen (1963) who found no relationship between attitude toward and achievement with programmed materials among a group of high school age subjects.

Finally, a more efficient teaching method might involve aspects of all three experimental methods used in this research. Programmed material and case studies used outside of class, either with overt or covert response, together with well-planned discussion sessions during class would combine the best elements of the three experimental approaches.

C. Recommendations for further study

1. A tentative model of the case study programmed instruction approach has been developed for this research. This model should be suitable for both evaluative and parametric type studies of this instructional approach. It is recommended that additional units in such topics as "The State and Education" and "Powers of School Boards" be developed so that future research may be based on a more extended exposure to the material.

2. A parametric study of the values of overt response and covert response in conjunction with discussion and no discussion conditions should prove useful in determining what combination of the response and discussion variables makes for most efficient learning.

3. The area of reading competency and its relationship to achievement and student satisfaction with programmed material deserves further exploration. Even at the graduate level there seems to be important differences among students in reading ability, speed, and comprehension; differences which probably affect learning via programmed material.

4. The technique of "programming" or planning discussion sessions so as to encourage optimum student response and a high level of reinforcement by the discussion leader should be evaluated experimentally. Discussion sessions with minimum verbal reinforcement by the leader, a higher level of reinforcement, and maximum reinforcement might be conducted and the effects on student achievement and attitude measured.

APPENDIX A

SAMPLE OF PROGRAMMED MATERIAL FROM

TOPIC E

Education, Religion, and the Constitution
(Suggested Time -- 55 Minutes)

In this section we will review briefly the historical background of two amendments to our Constitution under which questions regarding the relation of education and religion have arisen. We will study three court cases in this area and attempt to define legal theories that have been developed in the attempt to solve problems arising from the relation of religion, education, and the state.

Religion (denomination) (church)	1.	The religious wars of 15th and 16th century Europe, coupled with the growth of strong national states, caused many European kings to establish a particular____________ as the official state church.
"established" (state)	2.	These official or "____________" churches were supported by the civil authority and, in practically every instance, received financial assistance through some system of taxation.
established	3.	In its milder form, every citizen had to contribute to the support of the official state or____________ church whether he accepted its teachings or not.
participation	4.	In its more severe form, active ____________ in the worship services of the established church was required of all citizens.

financial (government official) participate in its services (or similar words)	C-5. An established church is at a minimum given ____________ support through taxes imposed on all citizens. In its more severe form all citizens are required to ** ____________ ____________
government (Congress)	6. Some delegates to the American Constitutional Convention opposed all forms of established churches; others wishes only to prevent the new national ____________ from ever setting up an official church.
YES - Incorrect. The application of the first amendment to the states did not occur until much later in our history. NO - Correct. In fact, at the time of the signing of the Constitution, 5 of the 13 original colonies still had their own established churches!!!	7. Thus, the first amendment reads in part: "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof;". Did this amendment, as originally worded above, restrict the power of the various states in establishing churches? YES ________ NO ________
establishing prohibiting (limiting) (restricting) states	8. The first amendment imposed two restrictions on Congress: it could not pass any laws either ____________ an official state church or ____________ the free exercise of any religion. The powers of the various ____________ were not restricted in any way by this amendment.

establishment exercise	9.	There are two clauses protecting religious freedoms in the first amendment. The first clause prohibits the ____________ of an official state church; the second clause prohibits any restriction on the free ____________ of religion.
establishment exercise	C-10.	Religious freedoms are protected both by the " ____________ clause" and the "free ____________ clause" of the first amendment.
establishment free exercise	11.	The ____________ clause prevents the government from setting up a state church, while the * ____________ clause prohibits government interference with individual religious practices.
first amendment sovereignty (power) (authority)	C-12.	The sovereignty of the various states in religious matters was not restricted by the * ____________ of the Constitution. However, later court interpretations of the fourteenth amendment served to restrict state ____________ in this matter.
states	13.	Many years after the adoption of the fourteenth amendment, the Supreme Court interpreted its clause "no state shall make or enforce any law which shall abridge the privileges and immunities of citizens of the United States" as making the first amendment apply also to the ____________.
restrict	14.	In this way, the fourteenth amendment, originally adopted primarily to protect the rights of Negro citizens eventually served to (restrict - enlarge) ____________ the sovereignty of the various states in other matters.

due process fourteenth child benefit	R-15.	The question raised previously in Cochran v. Louisiana was whether providing textbooks to parochial school children diverted public property to private use in violation of the * ______________ clause of the ______________ amendment. In its decision the Supreme Court proposed the * ______________ theory to justify this assistance. The next case involves clauses of both the fourteenth and first amendments. Please read Everson v. Board of Education, Case #7.
bus transportation costs parents of parochial school children (or similar answer)	C-16.	The question at issue in Everson v. Board of Education was whether the Board of Education could reimburse ** ______________ ______________ for ** ______________ ______________.
did not Everson	17.	The New Jersey courts ruled that the practice contested in the Everson case (did - did not) ______________ violate the state constitution. An appeal from this decision was then made to the United States Supreme Court by ______________ ______________.
due process establishment	18.	Everson maintained in his appeal that the Ewing Township practice deprived him of property in violation of the * ______________ clause of the fourteenth amendment, and, in addition assisted a particular church in violation of the ______________ clause of the first amendment.

due process fourteenth establishment first	C-19.	The Cochran v. Louisiana case involved an interpretation only of the * ________________ clause of the ________________ amendment, while the Everson case involved in addition the ________________ clause of the ________________ amendment.
states Constitution	20.	Since education is among the powers reserved to the various ________; the provisions of a state constitution and decisions of the highest state court in this area are final unless it can be shown that there has been a violation of a personal right protected by the Federal ________________.
state constitution	21.	Practically every state constitution contains provisions that prohibit public assistance to private schools. In this case, however, the New Jersey Supreme Court ruled that Ewing Township's reimbursement arrangement did not violate any provision of the * ________________.
United States Supreme Court	22.	Had the New Jersey Supreme Court ruled against the Board of Education, there would have been no basis for an appeal to the * ________________.
fourteenth first	23.	Since the New Jersey Supreme Court ruled against Everson, he could appeal to the United States Supreme Court on the grounds that certain personal rights protected by the ________________ and ________________ amendments of the Federal Constitution were being violated.

final right Constitution (Courts)	C-24.	The rule is that in cases involving education the decision of the state's highest court is __________ unless the case involves violation of a personal, protected by the Federal __________.
constitutional	25.	Just as Louisiana's practice of providing free textbooks to all was declared constitutional, so, also, was the New Jersey practice of providing free bus transportation declared __________.
child benefit	C-26.	In both Cochran v. Louisiana and Everson v. Board of Education, the Supreme Court developed its * __________ theory to justify expenditures supporting private education.
constitutional "of separation between church and state"	27.	Although in the Everson case the Supreme Court declared New Jersey's practice of providing free bus transportation for parochial school pupils as being __________, it interpreted the first amendment as, in Jefferson's words, building "a wall ** __________ __________.
Constitution (amendment)	28.	Jefferson's famous phrase regarding a "wall of separation" is not contained in the __________ itself. Jefferson originally made this statement in a letter written to certain New England Baptists. The Court, in attempting to interpret the Constitution, frequently reviews other writings of the founding fathers in a search for meaning.

free textbooks due process fourteenth amendment	R-29.	In Cochran v. Louisiana, the practice of a state in providing * ____________ was held not to violate the * ____________ clause of the * ____________ of the Constitution.
reimbursement for bus transportation costs of parochial school children establishment of religion first	C-30.	In Everson v. Board of Education, the practice of providing ** ____________ was held not to violate either the due process clause of the fourteenth amendment or the ** ____________ clause of the ____________ amendment.

APPENDIX B

SAMPLE CASE STUDY

Case Number 2

Helvering, Commissioner of Internal Revenue, et. al. V. Davis,

301 U.S. 619 (1937)

Background

Congress in 1935 passed legislation establishing a system of old age retirement benefits (the "Social Security" System) and providing revenue to support this system through two different types of tax; "an income tax on employees" and "an excise tax on employers."

In addition, certain features of the Act were designed to give financial assistance to states who established systems of unemployment compensation. This Act was immediately attacked in the courts on a variety of grounds. In another case decided during the same term (steward Machine Co. v. Davis, 301 U.S. 548) the Supreme Court ruled that:

> "the tax is not void as involving an unconstitutional attempt to coerce the States to adopt unemployment compensation legislation approved by the Federal Government... The problem of unemployment is national as well as local; and in promotion of the general welfare moneys of the Nation may be used to relieve the unemployed and their dependents in economic depressions and to guard against such disasters."

This case (Helvering v. Davis) arose when a shareholder of the Edison Electric Illuminating Company of Boston sought an injunction to restrain the Company from making the payments and deductions called for by the Act. The complainant, Mr. Davis, maintained that the Act, particularly Titles II and VIII, was unconstitutional in that it went beyond Congress' power to tax under the general welfare clause and that the Act interfered with rights reserved to the states by the 10th amendment.

His request for an injunction was denied by the district court but the Circuit Court of Appeals reversed the decision on the grounds that Title II of the Act was an invasion of powers reserved to the states or the people by the 10th amendment and that the tax upon employers was not an excise tax in the Constitutional sense.

The government then asked the Supreme Court to review the decision. The Court issued a writ of certiorari and, after reviewing the case, gave the following opinion: (excerpts)

> "The scheme of benefits created by the provisions of Title II is not in contravention of the limitations of the Tenth Amendment.
>
> Congress may spend money in aid of the general welfare... There have been great statesmen in our history who have stood for other views. We will not resurrect the contest. It is now settled by decision, United States v. Butler, supra. The conception of the spending power advocated by Hamilton and strongly reinforced by Story has prevailed over that of Madison, which has not been lacking in adherents. Yet difficulties are left when the power is conceded. The line must still be drawn between one welfare and another, between particular and general. Where this shall be placed cannot be known through a formula in advance of the event. There is a middle ground or certainly a penumbra in which discretion is at large. The discretion, however, is not confided to the courts. The discretion belongs to Congress, unless the choice is clearly wrong, a display of arbitrary power, not an exercise of judgment... nor is the concept of the general welfare static. Needs that were narrow or parochial a century ago may be interwoven in our day with the well-being of the nation. What is critical or urgent changes with the times...
>
> The problem is plainly national in area and dimensions. Moreover, laws of the separate states cannot deal with it effectively...

Whether wisdom or unwisdom resides in the scheme of benefits set forth in Title II, it is not for us to say. The answer to such inquiries must come from Congress, not the Courts. Our concern here, as often, is with power, not wisdom...

The decree of the Court of Appeals should be reversed and that of the District Court affirmed."

APPENDIX C

FINAL OPINION INSTRUMENT

SCHOOL LAW EXPERIMENT

STUDENT OPINION SURVEY

Name ____

This survey will take about 6 minutes to complete. Your opinions will be kept confidential and have absolutely no bearing on any aspect of the course in which you are now enrolled.

PART I

ATTITUDES TOWARD THE ______________________ METHOD OF INSTRUCTION.

DIRECTIONS: Each of the following statements represents a possible student attitude toward the method of instruction to which you have just been exposed. Please place a check mark under the heading which best reflects your attitude regarding each statement.

Everyone should answer this section.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1. I like this method of instruction very much.	___	___	___	___	___

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
2. I thought this method of instruction was a complete waste of time.	___	___	___	___	___
3. This method of instruction proved to be extremely boring and time consuming.	___	___	___	___	___
4. I would not object if this method of instruction were used in this class for the entire semester.	___	___	___	___	___
5. I would recomment this method to other teachers for use in their classes.	___	___	___	___	___
6. I would never use this method of instruction if I were teaching.	___	___	___	___	___
7. This method didn't arouse my interest very much.	___	___	___	___	___
8. This is one of the better methods of instruction to which I have been exposed.	___	___	___	___	___
9. This is the most interesting teaching method I have encountered in some time.	___	___	___	___	___

		Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
10.	I don't care for this method and I doubt if it appeals to other students.	___	___	___	___	___
11.	I feel that I learned little via this method of instruction.	___	___	___	___	___
12.	I would recommend this method of instruction for use in classes at other colleges and universities.	___	___	___	___	___
13.	This method of instruction should be used in more classes on all levels.	___	___	___	___	___
14.	This method of instruction proved to be quite distasteful.	___	___	___	___	___
15.	The monotony of this teaching method drives me to distraction.	___	___	___	___	___
16.	All professors should be encouraged to use this method for part of their instruction.	___	___	___	___	___

PART II

Opinions regarding the Case Studies used in this experiment

DIRECTIONS: For each statement below circle the letter preceding the item that best reflects your opinion. Add additional comments where necessary. Everyone should answer this section.

1. The case study material was:

   a. extremely interesting.

   b. rather dull.

   c. rather intersting.

   d. don't know or undecided.

   e. quite boring.

2. As far as the learning situation was concerned, the case studies:

   a. were of some value.

   b. don't know or undecided.

   c. didn't add or detract from what I learned.

   d. added a great deal to the learning situation.

   e. seemed completely unnecessary.

3. From the aspect of style and manner of expression, the case studies:

   a. were very well written.

   b. were rather poorly done.

   c. were above average.

   d. don't know or undecided.

   e. were below average.

4. As to length, the case studies:
   - a. were too long and detailed.
   - b. were just about the right length.
   - c. were too short and sketchy.
   - d. were not detailed enough.
   - e. don't know or undecided.

Add additional comments regarding the case studies here:

PART III

Opinions Regarding the Programmed Material used in this Experiment

DIRECTIONS: Circle the letter preceding the items that best reflect your opinion. Add your own comments where necessary. Only students who have used the program should answer this section.

1. This program:
   a. was an extremely interesting one to read.
   b. don't know or undecided.
   c. made rather interesting reading.
   d. was somewhat boring.
   e. was only of average interest.
2. The frames or steps in the program were:
   a. so small they were boring.
   b. too large for efficient learning.
   c. don't know or undecided.
   d. large enough to be challenging.
   e. just about right.
3. As to difficulty, the programmed material was:
   a. too easy.
   b. don't know or undecided.
   c. extremely difficult.
   d. rather difficult.
   e. just about right.

4. The format or physical arrangement of the programmed material:

   a. worked very efficiently.

   b. seemed O.K.

   c. don't know or undecided.

   d. caused some problems.

   e. seemed very inefficient.

5. The programmed material:

   a. added a great deal to the learning situation.

   b. was of questionable value.

   c. was of some value in the learning situation.

   d. detracted from the learning situation.

   e. don't know or undecided..

6. While using the programmed material, I missed the give and take of class discussion:

   a. very much.

   b. not at all.

   c. to a minor degree.

   d. don't know or undecided.

   e. to some extent.

7. I began to feel fatigued after using this program for approximately:

   a. 50 minutes.

   b. 100 minutes.

   c. 150 minutes.

   d. 200 minutes.

   e. had no feeling of fatigue.

   f. don't know or undecided.

Additional comments about the program:

PART IV

Opinions regarding Potential Uses of this programmed material

DIRECTIONS: Circle the letter preceding the items that best reflect your opinion. Add your own comments where necessary. Only students who have used the program should answer this section.

1. This program:
   - a. seems quite suitable for use by most graduate students.
   - b. should never be used on the graduate level.
   - c. don't know or undecided.
   - d. should be used by all graduate students at some point in their course of study.
   - e. seems suitable for only a few graduate students.
2. This material seems best suited to:
   - a. continual use during class time.
   - b. occasional use during class time.
   - c. infrequent use during class time.
   - d. don't know or undecided.
   - e. should never be used during class time.

3. For individual study, I believe this type of program is:

   a. definitely superior to a conventional text.

   b. don't know or undecided.

   c. better than a conventional text.

   d. inferior to a conventional text.

   e. about the same as a conventional text.

4. I would recommend that material of this type:

   a. be developed for other graduate courses.

   b. be developed for both undergraduate students and graduate courses generally.

   c. be confined to use with undergraduate students.

   d. don't know or undecided.

   e. be restricted to use only on the elementary or secondary level.

Please give additional comments or recommendations regarding possible uses of this material:

PART V

Opinions regarding the Discussion Approach

DIRECTIONS: Circle the letter preceding the item that best reflects your opinion. Add your own comments where necessary. Only those in the discussion group should answer this section.

1. The discussion guides were:
   - a. very well organized.
   - b. well organized.
   - c. about average.
   - d. very poor.
   - e. don't know or undecided.
2. The discussion guides:
   - a. didn't add much to the learning situation.
   - b. were of some help.
   - c. were unnecessary.
   - d. don't know or undecided.
   - e. were very helpful.
3. The discussions themselves were:
   - a. very stimulating.
   - b. rather boring.
   - c. about average.
   - d. don't know or undecided.
   - e. above average in stimulating thought.

4. I thought that group participation in the discussion was:
    - a. about average.
    - b. well above average.
    - c. very poor.
    - d. don't know or undecided.
    - e. below what would be normally expected.

Additional comments regarding the discussion approach or any other aspect of the experiment:

Please mail this survey in the return envelope as soon as possible.

Thank you for your help.

APPENDIX D

SAMPLE ITEMS FROM THE
TESTING INSTRUMENT

Name________________________________Date__________

DIRECTIONS: For the multiple choice items below, circle the letter preceding the item you would select as the best or most complete answer.

1. In practice, the American Constitution consists of:

a. The written document and its amendments.

b. The written document, amendments and unwritten practice.

c. The written document and amendments as interpreted by government officials.

d. Both b and c.

e. Both a and c.

4. Article I, Section 10, of the Constitution:

a. Restricts state power regarding ex-post-facto laws and contracts.

b. Enlarges state power regarding ex-post-facto laws and contracts.

c. Restricts both state and federal powers in the two areas mentioned above.

d. Restricts state and federal power regarding contracts only.

8. A citizen who had been denied due process of law by a Federal agency would appeal to the courts under provisions of:

a. The 14th Amendment.

b. The 14th and 1st Amendment.

c. The 1st and 5th Amendment.

d. The 5th Amendment.

12. The general welfare clause usually referred to in court case is found in:

a. The Preamble

b. Article I, Section 8.

c. Article I, Section 10.

d. Both a and b.

e. Both a and c.

16. In carrying our their governmental responsibilities, the Constitution is interpreted by officers of:

a. The Executive Branch.

b. The Legislative Branch.

c. The Judicial Branch.

d. All of the above.

e. None of the above.

20. Helvering v. Davis is an important case for education because:

a. The Social Security Act was held to be constitutional.

b. Congress' power to tax for social welfare legislation was upheld.

c. The general welfare clause was interpreted broadly as giving Congress power to assist the states and secure their cooperation in projects of national concern.

d. Federal aid to education was not questioned.

24. Police power refers to the right of the state:

a. To control lawbreakers and criminals.

b. To protect the indiviual from group pressures.

c. To overrule constitutional guarantees of freedom.

d. To place restrictions on individuals or their property in the interest of the common good.

28. Certain expenditures of tax moneys to assist parochial schools have been declared constitutional under the Supreme Court's:

a. State Benefit theory.

b. Child benefit theory.

c. Representative taxation theory.

d. Public welfare theory.

32. As originally enacted, the 1st Amendment restricted the power of:

a. Both the Federal and state governments.

b. Both Congress and the President.

c. Congress.

d. All of the above.

36. The decision in Everson v. Board of Education involved interpretations of:

- a. The due process clause of the 5th Amendment and the establishment clause of the 1st amendment.
- b. The due process clause of the 14th Amendment and the establishment clause of the 1st Amendment.
- c. The due process clause of the 14th Amendment and the free exercise clause of the 1st Amendment.
- d. The due process clause of the 5th Amendment and the free exercise clause of the 1st Amendment.

40. The Supreme Court first adopted the wall of separation doctrine regarding the 1st Amendment in:

- a. Pierce v. Society of Sisters.
- b. Illinois ex rel v. McCollum.
- c. Everson v. Board of Education.
- d. Cochran v. Louisiana

APPENDIX E

DISCUSSION GUIDE

School Law

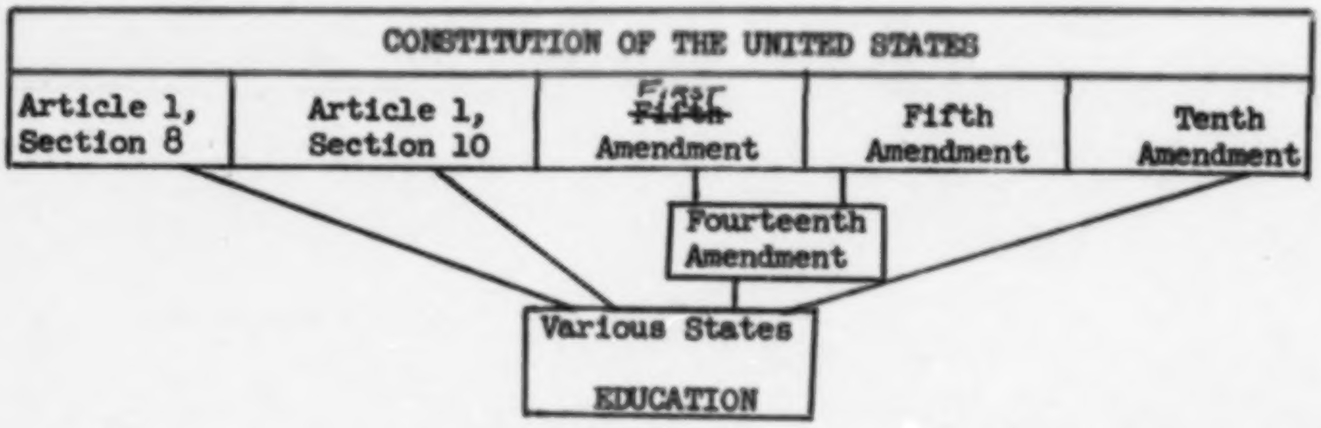
Topic - Federal Relations to Education

Discussion Guide

I. Education and the Constitution

- A. The nature of a constitution.
    1. What is a constitution?
    2. What is its purpose in government?
    3. How do constitutions differ from country to country?
- B. Our Federal Constitution and Education
    1. The place of education in our Constitution.
    2. Sections of the Constitution pertaining to education.

       (See Chart)


CONSTITUTION OF THE UNITED STATES
Article 1, Section 8
Article 1, Section 10
First ~~Fifth~~ Amendment
Fifth Amendment
Tenth Amendment
Fourteenth Amendment
Various States
EDUCATION

a. Article 1, Section 8 - the general welfare clause.

b. Article 1, Section 10 - restrictions on contracts and ex-post-facto laws.

- c. The First Amendment - Freedom of speech and religion.
- d. The Fifth Amendment - Judicial rights and due process. (Federal)
- e. The Tenth Amendment - Reserved powers.
- f. The Fourteenth Amendment - Equal protection and due process. (State)

II. Education and the General Welfare

- A. Interpreting the general welfare clause.
    - 1. Hamilton's viewpoint.
    - 2. Madison's viewpoint.
    - 3. The Supreme Court's role in interpreting the Constitution.
    - 4. The role of other governmental branches in interpretation.
- B. U.S. v. Butler.
    - 1. Facts of the case.
    - 2. The court's reasoning and decision.
    - 3. Importance to education.
    - 4. Implications for Federal regulation of education.
- C. Helvering v. Davis.
    - 1. Facts of the case.
    - 2. Reasoning and decision of the Court.
    - 3. Importance of this case to education.

III. Contracts, Education and the Constitution.

- A. Article 1, Section 10 of the Constitution and the education function of the various states.
- B. Dartmouth College v. Woodward.
    - 1. Facts of the case.
    - 2. Reasoning and decision of the court.
    - 3. Implications of this case for education.

IV. The Fourteenth Amendment and the Police Power

- A. The nature of the police power.
- B. Due process of law and exercises of the police power.
    - 1. Role of the Supreme Court.
- C. Meyer v. Nebraska.
    - 1. Facts of the case.
    - 2. Reasoning and decision of the Court.
    - 3. Importance to education.
- D. Pierce v. Society of Sisters of the Holy Name of Jesus and Mary
    - 1. Facts of the case.
    - 2. Reasoning and decision of the Court.
    - 3. Implications for education.
- E. Cochran v. Louisiana.
    - 1. Facts of the case.
    - 2. Reasoning and decision of the Court.
    - 3. Importance to education.

V. Education, Religion and the Constitution.

- A. Church and state in colonial America.
    - 1. The nature of an established church.
    - 2. Historical background of the First Amendment.
    - 3. Restrictions imposed by the First Amendment.
- B. Everson v. Board of Education.
    - 1. Facts of the case.
        - a. Role of the state courts.
    - 2. Reasoning and decision of the Court.
        - a. The child benefit theory.
    - 3. Implications for education.

C. **Illinois ex rel v. McCollum.**

1. **Facts of the case.**
2. **Reasoning and decision of the Court.**
3. **Implications for education.**

D. **Zorach v. Clauson.**

1. **Facts of the case.**
2. **Reasoning and decision of the Court.**
3. **Comparisons with McCollum case.**
4. **Implications for education.**

VI. **Summary and Review**

BIBLIOGRAPHY

Beck, Carleton E., "Teacher Education Self-Taught," *Journal of Teacher Education*, Vol. XIV, No. 1, March, 1963.

Edwards, Allen L., *Experimental Design in Psychological Research*. Holt, Rinehart and Wilson, New York, 1960.

Edwards, Allen L., *Statistical Methods for the Behavioral Sciences*. New York: Rinehart and Company, 1958.

Eigen, Lewis D., "High School Student Reactions to Programmed Instruction," *Phi Delta Kappan*, Vol. XLIV, No. 6, March, 1963.

Evans, J. L., "An Investigation of 'Teaching Machine' Variables Using Learning Programs in Symbolic Logic." Ph.D. Dissertation, University of Pittsburgh, 1960.

Garrett, Henry B., *Statistics in Psychology and Education*. New York; Longmans, Green and Company, 1953.

Goldbeck, R. A., "The Effect of Response Mode and Learning Material Difficulty on Automated Instruction." Technical Report No. 1. American Institute for Research, Pittsburgh, Pa., 1960.

Goode, William J. and Hatt, Paul K., *Methods in Social Research*. McGraw Hill Book Company, New York, 1952.

Pressey, Sidney, "A Simple Apparatus Which Gives Tests and Scores -- and Teaches." *School and Society*, XXIII. (March, 1926.)

Schramm, Wilbur, *Programmed Instruction Today and Tomorrow*. (New York: Fund for the Advancement of Education, 1962.)

Silberman, R. E., and Alter, M., "Response Mode, Pacing and Motivational Effects in Teaching Machines." *Programmed Instruction I* (October, 1961.)

Skinner, B. F., "The Science of Learning and the Art of Teaching." *Harvard Educational Review*, XXIV (1954.)

Spurlock, Clark, *Education and the Supreme Court*. University of Illinois Press, Urbana, Illinois (1955.)

Walker, C. C. and Stolurow, L. M., "A Comparison of Overt and Covert Response in Programmed Learning." University of Illinois, 1962. (42 page mimeographed report.)

Winer, Ben J., *Statistical Principles in Experimental Design*. McGraw Hill, New York, 1962.